	BEARCH REQUEST F	OFM 🕏 💉	9/26/07
Requestor's	Senal		
Name VIIIV //	Airdia Sectal Numb	өг <u>эх ОЭ-556</u> -	<u> </u>
Bulles 7. 85 CO	Phone: 605.121	Antiunity de	<u>\$3</u>
Pleasewrite a detailed statement of search that may have a special meaning. Give ex- that copy for the sequence. You may includ	topic: Describe specifically as possible amples or relevant cuations; authors key	the subject matter to be scarched words, etc., it known ! For seque	Define any terms
	a copy of the organistical and/or most re-	(evant claim(s)	
Please search			
2.00			75-199
Dissig 1, aa 1-24	plus aa 200-14	04 (delete al	
Z AAS			
1 1300, an 1-25	plus 67-1404	(delete as 26	-66)
501 aa 1-25 ph	is aa 67-106 plus	aa 200-1404	
) Sagl, aa 1-25 pl	(delete aa 26-61	6 and 107-199)	
Daa 200:-1140 d	SeglV		
8). Secil ad 25-	end (delete aa	1-34)	
6) See 1 aa 1-66 7) 800	ölus 105-end (detete aa 67-10	4)
7) 500	Tus as 200-1404 (West on 157-199)
Seal an 1-106 #	lus aa 200-1404 (d	lilete aa 107 - 1998	
) Seg. 1 as 200-11671	tus as 200-1404 (d tus as 200-1404 (d JYJJ) Sig 1 (ds. 2)	∞-1212-	
) Segl ac 200-1263		Point of Contac Toby Port	t.
		Technical Info Spe CM16A04 703-308-3534	

1/29	STAFFUSE ONLY		
Daie completed 4/29 Searcher Love + B. OBryon	Search Site	Vendors	
			G Stitle
	EM-1	S	TN:
Hapsed time: 20	Pre-S		nalog
Terminal time: //26 Blessed time: ZO PU time: Total time:			
Resed time: ZO	Pre-S Type of Search	Tuence G	nalog PS

CONT. AND SON OF THE S

				1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -		(A.C.)	
			Fig.		4		
	·						
		•					
					as		
•							

```
RA Carden M.J.;

RA Carden M.J.;

RL SUBMITTED THE EMBL/GenBank/DDBJ databases.

C SUBMITTED THE TWO SUBLIX CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

CC SUBSERVED BY THE TWO SMALLER N PROTEINS.

CC SUBSERVED BY THE TWO SMALLER N PROTEINS IN THE FORMATION OF CC INTERFILMENT CALIBER.

CC SAXONAL CALIBER.

CC SAXONAL CALIBER.

CC STREIL LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE COINCIDENT WITH A CHANGE IN THE NUTREMEDIATE FILAMENT FUNCTION.

CC STANLLARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC SATURE NEW SAGURNEE DIFFERS FROM THA SHOWN IN POSITIONS 534

CC SATURE AND IS SHORTER DUE TO FRAMESHIFTS.

CC SATURE AND IS SHORTER DUE TO FRAMESHIFTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                   Ή.
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLU'RICH (ACIDIC).
50 X 6 AA TANDEM REPEATS OF K-S-P-A-E-A.
GLU/LYS-RICH.
                                                                                                               Sequence FROM N.A.
MEDINE-8908913145094;
MEDINE-8908913145094;
Shneidman P.S., Carden M.J., Lees J.F., Lazzarini R.A.;
The structure of the largest murine neurofilament protein (NF-H)
revealed by CDNA and genomic sequences.";
Brain Res. 464:217-231(1988).
  Julien J.-P., Cote F., Beaudet L., Sidky M., Flavell D., Grosveld
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M24496; AAA39813.1; --
REMBL; M2349; AAA39813.1; --
REMBL; M24494; AAA39813.1; JOINED.
REMBL; M35131; AAA39813.1; JOINED.
REMBL; M35131; AAA39809.1; ALT_FRAME.
REMBL; M35131; AAA39809.1; ALT_FRAME.
R PIR; A43778; A43778.
R PIR; A43778; A43778.
R PIR; A43778; A13778.
R PIR; A43778.
R PI
                                               "Sequence and structure of the mouse gene coding for the largest neurofilament subunit.";
Gene 68:307-314(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COIL 2A.
LINKER 2.
COIL 2B.
K -> QA (IN REF. 2 AND 3).
A -> AR (IN REF. 2 AND 3).
S -> T (IN REF. 2 AND 3).
L -> G (IN REF. 2 AND 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COIL 1B.
LINKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COIL 1A.
LINKER 1.
                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
STRAIN-SWISS WEBSTER, TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     517
886
1087
129
141
239
261
261
287
408
1133
1199
492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98
4409
4409
1130
1142
240
284
288
288
1133
1133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
```

```
43;
                                                                                                                                                                                                                                                                                                                                                                        105 KKTKKVIESE--EITEEHSVS-ENQESSSSSSSSSSTIWKIKSSKNSAANRELQKKLK 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            208 VSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTNKQTSTDGKEKTT 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     383 APTITKSAPTIPKEPA----PITPKKPA----PITPKEPA----PTIPKEP----TPTTP 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           479 TIPKEPA----PITIKSAPITIKEPAPITIKSAPITPKE-PSPITIKEPA----PITPKE 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       698 TTTKEPTSTTSD--KPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTT 755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTT 875
                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 VK--------DNKKNRTKKKPTPKPPVDEAGSGLDNGDFKVTTPDTSTTQHNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433 VKSEEMIKVVE------KSEKETVIVEGQTEEIRVTEGVTEEEDKEAQGQEGEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  : : : | | :| : | | : | : | | : : | | : : | | : : | | | : : | | | : : | | | : : | | | : : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIPTIKSAPTIPKEPAPTITKSAPITPKEP-APITIKEPAPTIPKEPAPTITKEP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                651 KSPA---EAKSPAEAKSPAEAKSPAEVKSPGEAKSPAEPKSPAEA--KSPA----EVKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEPAPITKEPAPITPKEPA----PTAPKKP----APITPKEPAPITFKEPAPITKEPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIEVKSPEKAKTPVKEGAKSPAEAKSPEKAKSPVKEDIKPPAEAKSPEKA-KSPVKEGAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   590 TPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEE-PAPTTPKAAAPNTPKEPAPTT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --PKEPAPTTP-----KEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      756 PKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 7.6%; Score 551; DB 1; Length 1087; Similarity 26.8%; Pred. No. 5.8e-15; 4; Conservative 96; Mismatches 358; Indels 186;
  P -> PREAKSP (IN REF. 3).
MISSING (IN REF. 3).
G -> A (IN REF. 3 AN).
V -> M (IN REF. 2 AND 3).
T -> N (IN REF. 2 AND 3).
                                                                                                                                                116612 MW; 57BAC76A38ED1CB9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1060 AEKSSSTDQKESQPPE-----KTTEDKATKGEK 1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            876 TAKDKTTERDLRTTPETTTAAPKMTKETATTTEK 909
                                                                                                                                                                                                                                                                               Matches 234; Conservative
     551 55
689 71
714 71
814 81
843 84
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1018
  FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                         Op
                                                                                                                                                                                                                                                                                δŽ
                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                          qq
                                                                                                                                                                                   Qγ
                                                                                                                                                                                                         g
                                                                                                                                                                                                                                 á
FT
FT
FT
SO
                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profilt institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MAIN ROOT.

BUDGELLUAR LOCATION: EXTRACELLULAR MATRIX.

STHE LOCATION: EXTRACTERISTIC REPEAT OF THE PENTAPEPTIDE STH. EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EXTENSIN.
H.A-P-P.
H.A-P-P.
Z X 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.
                                                                                                                                                                           PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKET--A 789
                                                                                                                                                                                                                          677
                                      757
                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    [1]
STRAIN=CV. XANTHI, TISSUE-Leaf;
STRAIN=CV. XANTHI, TISSUE-Leaf;
MEDLINE=90128263; PubMed=2612909;
MEDLINE=90128263; C.J.;
Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation.";
glycoprotein gene in lateral root initiation.";
THE MECHANICAL PRECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE
            EPSPTTTKEPAPTTPKEPAPTTPKKPA----PTTPKEPAPTTPKEPAPTTKKPAPTAP
                               KEPAPITPKETAPI-----TPKKLIPTIPEKLA-----PITPEKPAPITPEELAPITPEE
                                                                           PTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPT
                                                                                                                             KPTISPEKPTISTEK------PIPTE-KPTIPTE-----ETTISTEKLTIPT
                                                                                                                                                           TLKEPAPTTPKKPAPKELAPTTTKEPTSTT-----SDKPAPTTPKGTAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                              EXTENT TOBAC STANDARD; PRT; 620 AA.
P13983;
01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 21, Last sequence update)
01-MAR-1992 (Rel. 21, Last annotation update)
EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; S06733; S06733.
Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
                                                                                                                                                                                                                                                                                                                                                                                                                           HRGPNT3.
Nicotiana tabacum (Common tobacco).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X13885; CAA32090.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             73
73
151
242
                                                                                                                                                                                                                                                                                        PASCKSPRPS 976
                                                                                                                                                                                                                                                                  PTTPKEPAPT 799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLYCOSYLATED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      70
148
229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydroxylation.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                        EXTN_TOBAC
                                                                                                                                            815
                      513
                                                                    568
                                                                                                                    618
                                                                                                                                                                                                                    732
                                                                                                                                                                                                                                                                  190
                                                                                                                                                                                                                                                                                         196
                                                                                                                                                                                          928
                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                          Op
                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                   οy
                                                                                                                   δ
                                                                                                                                          рp
                                                                                                                                                                   δy
                                             g
                                                                      ò
                                                                                          g
```

```
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        727
                                                                                                                                                                                       297 AETITK-GPALITP--KEPIPIT----PKEPASITPKE---PIPTIKSAPITPKEPAP 345
                                                                                                                                                       Gaps
                                                                                                                                                                                                            24 AEATTQYGGYLPPPVTSQPPPSSIGLSPPSAPTTTPPSRGHVPSP---RHAPPRHAYPPP 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            728 APTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PPPSPIYSPPPQVQPLPPTFSPPPRRIHLPPPPHRQP---RPPTPTYGQPPSPTFSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               539 PPPRQIHSPPPPHWQPRTPTYGQPPSPPTFSAPPPRQIHSPPPHRQPRPPTPTYGQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TAPTTPKEPAPT -- TPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-----LPPTYSPPPAYSPPPPTYSPPPPTYSPPPAYAQPPPPPPTYSPPPAYSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   611 APITPEEPTPTTPEEPAPTTPKAAAP----NTPKEPAPTTPKEPAPTTPKETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTTKSAPTTPKEP-----APTTTKEPAPT---TPKEPAPTTTKEPAPTTTKSAPTTPKE
                                                                                                                                                                                                                                                                                            SHGHLPPSVGGPPPHRCHLPPSRGFNPPSSPVISPSHPPPSYGAPPPSHGPGHLPSHGQR
                                                                                                                                                                                                                                                                                                                                  PAPTTPKKPAPT - - - TPKEPAPTTPKEPAPTTK - - EPAPTTPKEPAPTAPKK
                                                                                                                                                                                                                                                                                                                                                      ---APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPT----TPKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    250 RHLPPSPRRQPQPPTYSPPPAXAQSPQPSPTYSPPPPTYSPPPSPTYSPPPAXSPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            561 KPAPT---APKEPAPTTPKETAPTTPKKL-TPTT----PEKLAPTTPEKPAPTTPEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                     452 PAPTTPKEPAPTTPKEPAPTTTKEPSPTTPK---EPAPTTTKSAPTTTKEPAPTTTKS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NFH_MOUSE STANDARD; PRT; 1087 AA.

NFH_MOUSE STANDARD; PRT; 1087 AA.

1019.04 (Rel. 06.195);

101-NOV-1990 (Rel. 23, Last sequence update)

20-AUG-1992 (Rel. 40, Last annotation update)

20-AUG-101 (Rel. 40, Last annotation update)

NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT PROTEIN)

NEFH OR NFH.
                                                                                                                                                       88;
1.
CONTAINS THE SER-PRO(4) REPEATS.
3 X APPROXIMATE TANDEM REPEATS.
4; 641DD2278AB28524 CRC64;
                                                                                                                     Length 620;
                                                                                                                                                         304; Indels
                                                                                                                    7.6%; Score 551; DB 1;
27.2%; Pred. No. 3.6e-15;
live 64; Mismatches 304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PTPKALENSPKEPGVPTTKTPAATKP 871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE=89121513; PubMed=3220257;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | : |: ||
| : |-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Rodentia;
                                                                       MM:
 235
242
620
600
65406 P
                                                                                                                               Query Match
Best Local Similarity 27.2'
Matches 170; Conservative
 229
236
205
499
620 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |:|
PSP-
                                                        DOMAIN
SEQUENCE
   REPEAT
REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NFH_MOUSE
                                                                                                                                                                                                                                                                      346
                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                             141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507
                                                                                                                                                                                                                                                                                                                                           397
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                              959 IMNKPEETAKPKDRATNSKATTPKPQKPTKAPKK---PTSTKKPKTMPRVRKPKTTPTP- 1014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -I- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
                                                                                  958
                                                                                                                                                                   747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE ZONA PELLUCIDA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Large-scale sequencing of two regions in human chromosome 7q22; analysis of 650 kb of genomic sequence around the EPO and CUTLI loci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Testis;
MEDLINE-97271566; PubMed-9126492;
Gao Z., Harumi T., Garbers D.L.;
"Chromosome Localization of the mouse zonadhesin gene and the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS, THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDICTAL ISTHMUS. DOMAIN: THE VWFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2). SIMILARITY: CONTAINS 3.5 MAM DOMAINS.
850 ALENSPK----EPGVPTT--KTPA----ATKPEMTTTAKDKTTERDLRTTPETTTAAPK
                                                                       899 MIKETATITEKITESKITATITQVISTITQDITPFKITILKITTLAPKVITTKKIIITTE
                                    689 AKPINPSIYKAKPSYPPIYKAKPSYPPIYKAKPSYPPIYKAKPIYKAKPIYPSIYKAKP-
                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1-2379 FROM N.A.
MEDLINE-99018118; PubMed-9799793;
Glockner G., Scherer S., Schattevoy R., Boright A., Weber J.,
Tsui L.C., Rosenthal A.;
                                                                                                                                                                                                                                                                                                                                 JUMAN
STANCHOLAN
STANCHOLAN
OGY493; O00218;
20-AUG-2001 (Rel. 40, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                   1015 ----- RKMTSTMPELNPTSR 1029
                                                                                                                                                                                                                                                        849 YNKKISYPSSYKAKTSYPPAYKPTNR 874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 2338-2700 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  reveals 17 genes.";
Genome Res. 8:1060-1073(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 zonadhesin gene (ZAN).";
Genomics 41:119-122(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIGNALING.
                                                                                                                                                                                                                                                                                                                             ZAN_HUMAN
                                                                                                                                                                                                                                                                                                                                                a
                                                                         ö
                                                                                                       g
                                                                                                                                              ð
                                                                                                                                                                               q
                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                      g
```

```
219 PINPRPSLPPNSDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTTSAKETQSIEKT 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        445 PVKVLPELPPVSPVSS----TGPSETTGLTENPTISTK------KPTVSIEKP 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPT-----PTTI 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       334 KSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA-PTTTKSAPT 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    648 -PITPIE-KPIIPIEKPIISTEKPIIPIEK-PIISPEKLIIPIEKLIIPT---EKPIIPT 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 107; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ТРКЕРАРТТРККРАРТТРКЕРАРТТРКЕРТТРКЕРАРТТКЕРАРТТРКЕРАРТТРКЕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --EKPSIPTEK---PTISMEETIISTEKPTICPEKPTIPTEK---PTIPTEKSTISPEK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          453 APTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 2700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NNOKMA -> RAGPGP (IN REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           293013 MW; 80E60CC0B12277B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Cell adhesion; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MUCIN-LIKE DOMAIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.6%; Score 555.5; DB 1;
32.6%; Pred. No. 8.3e-15;
ive 64; Mismatches 240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGF-LIKE.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (PARTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                   P_rich_extensn.
TIL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VWFD
VWFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VWFD
                                                                                                                                                                                 Pfam: PF00629; MAM: 3.
Pfam: PF01826; TIL; 4.
Pfam: PF02345; TILe; 4.
Pfam: PF00345; vwd; 4.
PRINTS; PR01217; PRICHEXTENSN.
                                              EGF-like.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Glycoprotein; Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                            e. ⊢. 4.
                                                                                                                          InterPro; IPR003328; TILa.
InterPro; IPR001007; VWFC.
                                                                                                                                                                                                                                                                                         SMART; SM00137; MAM; 2.
SMART; SM00214; VWC; 1.
SMART; SM00011; VWC_def;
SMART; SM00216; VWD; 3.
                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS01186; EGF_2;
PROSITE; PS00740; MAM_1;
PROSITE; PS50060; MAM_2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           403
1023
1099
1618
1737
1832
1832
2136
2505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1861
2292
2684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 199; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <1
112
161
322
483</pre>
                                                                                                       InterPro;
InterPro;
                                              interPro;
                                                                                   interPro;
                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CONFLICT
                                                                                                                                                                    InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  488
DOR NO DO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

32;

EMBL; AF053356; AAC78790.1; -.

g δλ g ò Pp δ g ò

δ g ŏ

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                  PQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPN 1042
                                                                                                                                                                                                                                                                                                                                                                                               QT-PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPML 1101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       hydroxyproline-containing decapeptide in the adhesive protein of the mussel, Mythius edulis L. ".

J. Biol. Chem. 258:2911-2915(1983).

J. Biol. Chem. 258:2911-2915(1983).

J. Biol. Chem. 258:2911-2915(1983).

J. Brouce One OF THE STROMEST WARPER INDOLUBLE GIVE. THE MUSSEL'S ADHESIVE IS A BUNDLE OF THERADS, CALLED A BYSSUS, FORMED BY A FIBROWS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

J. SUBCELLUAR LOCATION: SECRETED.

J. TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

J. TISSUE SPECIFICITY: COMPOSED OF REPEATS OF A DECAPERTIDE.

J. PIN: THE DECAPEPTIDE A. F. P. P. P. T. Y. KIS POST TRANSATIONALLY MODIFIED AS FOLLOWING THE SIXTH AND SEVEWIH RESIDUES ARE HYDROXYLATED AND THE PENULTIMATE IS A 3,4-DIHYDROXYPHENYLALANINE
                                                 701 APDIRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAP-----DIRPAPGSTAP-- 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. MEDIANG-1367451; MEDLINE-91025829; PubMed-1367451; Filpula D.R., Lee S.M., Link R.P., Strausberg S.L., Strausberg R.L.; "Structural and functional repetition in a marine mussel adhesive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30-MAY-2000 (Rel. 39, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ADHESIVE PLAQUE MATRIX PROTEIN (POLYPHENOLIC ADHESIVE PROTEIN) (FOOT
  ---IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEM
                                                                                                  TTTAKDKTTERDLRTTPETT - - - - - TAAPKMTKETATTTEKTTESKITATTTQVTSTT
                                                                                                                               PGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPA
                                                                                                                                                                                                                                                                                                                                                P--GSTAPPAHGVTSAPDTRP---APGSTAPPAHGVTSAPDTRPA------PG
                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GVTSAPDNRPALGSTA----PPV
                                                                                                                                                                                                  927 TQDTTP--FKITTLKTTTLAPKVTTTKKTITTT--EIMNKPEETAKPKDRAINSKATTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mytilus edulis (Blue mussel).
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
Mytiloidea; Mytilidae; Mytilus.
NCBI_TaxID=6550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AND POST-TRANSLATIONAL MODIFICATIONS. Pubmed=6298211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SDEINICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPS 1142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      875 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      910 STAPPAHGVTSAPDTRPAPGSTAPPAH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Biotechnol. Prog. 6:171-177(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (DOPA) DERIVED FROM TYROSINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequ
20-AUG-2001 (Rel. 40, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEIN 1) (MEFP1) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PARTIAL SEQUENCE,
MEDLINE=83135732;
       TPTTTKEPTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Waite J.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FP1_MYTED
Q25460;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein.
                                                                                                                                                                                                                                                                                                                                                                                                            1043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1102
                                                                                                                                                        754
                                                                                                                                                                                                                                                                                                          983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FP1_MYTED
```

Physical Control of the Control of t

```
26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TANDEM REPEATS OF Y-K-[PA]-K-[LP]-[ST]-Y-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                424
                                                                                                                                                                                                                                                                                                                                  HNPTSPPSSKKAPPPSGASQTIKSTIKRSPKPPNKKKTKKVIESEEITEEHSVSENQESS 129
                                                                                                                                                                                                                                                                                                                                                                                                                         SSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLD 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   414
                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                               35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYPSTYKAKPSYP----PTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          579 APTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KTTYPPTYK--------PKISYPPTYKAKPSYPATYKAKPSYPPTYKAKPSYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 746 TILKEPA--PTTPKKPA--PKELAPTTTKG-PTSTTSDKPAPTTPKETAPTTPKEPA--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KELAPTTTKEPTSTTSDKPAPTTPKGTAPTT----PKEPAPTTPKEPAPTTPKGT---AP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTPKKPA-PTT----PETPPPTTSEVSTPTTKE----PTTIHKSPDESTPELSAEPTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKEPAPTTPKEPAPTTPKE----PAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----KPKMTYPPTYKPKPSYPP-----TYKSKPTYKPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---YPPTYKAKPSYPSSYKPKKTYPPTYK-----PKLTY----PPTYKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                310 KEPTPTT----PKEPASTTPKEPTPTTKSAPTTP----KEPAPTTKSAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITKEPA-PITPKEPAPITIKEPAPITIKSAPTIPKEPA--PITPKKPA--PITPKEPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYKAKPTYKAKPTYPSTYKAKPSYPPTYKAKPTYKAKPSYPPTYKAKPSYPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       - PTTPKEPT - PTTPKEPAPTTKEPAPTTPKEPAPTAPKKPA - - PTTPKEPA - - PTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           260 PPTYKAKPTYKAKPTYKAKPTYKAKPSYP----PTYKAKPSYPPTYKAKPSYPPTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A - - PTTTKEPS - - PTTPKEPA - PTTTKSAPTTTKEPA - PTTTKSAPTTPKEPSPTTTKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          523 APTTPKEPAPTTPKKPA--PITPKEPA--PTTPKEPAPTTTKKPAPTAPKEPAPTTPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SYPPTYKAKPIYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKP------T
                                                                                                                                                                                                                                                                                                                                                                               -----YKTSYSAPYKPPTYQPLKKKVD------YRPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSYPPTYGSKTNYLPLAKKLSS-----YKPIKTTYNAK-----TNYPPVY-----
                                                                                                                                                                                                                                                                                               260;
                                                                                                                                                                                                                                                  Length 875;
                                                                                                                                                                                                                                                  Query Match
7.7%; Score 559; DB 1; Length 87
Best Local Similarity 26.9%; Pred. No. 2.4e-15;
Matches 281; Conservative 102; Mismatches 403; Indels
                                                                                                                                                                         P-P-[ST].
; 6EA85312748CAACE CRC64;
               InterPro; IPR002964; Adhesive_plaq.
InterPro; IPR002965; P_rich_extensn.
PRINTS; PR01216; ADHESIVEI.
PRINTS; PR01217; PRICHEXTENSN.
                                                                                                                                                                                           100412 MW;
EMBL; X54422; CAA38294.1;
                                                                                                                                                  870
                                                                                                             Repeat; Hydroxylation.
                                                                                                                                                                                           875 AA;
                                                                                                                                                  29
                                                                                                                                                                                                                                                                                                                                                                                      3 HEPVYKP--
                                                                                                                                                                                             SEQUENCE
                                                                                                                              NON_TER
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      693
                                                                                                                                                                                                                                                                                                                                              20
                                                                                                                                                                                                                                                                                                                                                                                                                                  130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9/
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                             IS ALSO PRODUCED.
-I- ALTERNATIVE PRODUCTS: VARIOUS VARIANTS ARE PRODUCED BY ALTERNATIVE
                                                                                                                                                                                                                                                                          TUMORS, SUCH AS BREAST CANCER.
                                                                                                                                                                                                                                                                                                            POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC. IT VARIES FROM 21 TO 125 IN THE NORTHERN EUROPEAN POPULATION. THE MOST FREQUENT ALLELES CONTAINS 41 AND 85 REPEATS. SIMILARITY: CONTAINS 1 SEA DOMAIN.
                                                                                                                                                                                                      CYTOSKELETON.
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. A SECRETED FORM
"preoperative diagnosis of thyroid papillary carcinoma by reverse transcriptase polymerase chain reaction of the MUC1 gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1- TISSUE SPECIFICITY: ABERRANTLY EXPRESSED IN HUMAN EPITHELIAL
                                                      TISSUE-Lung;
MEDLINE-96181716; PubMed=8604237;
Yu C.J., Yang P.C., Shew J.Y., Hong T.M., Yang S.C., Lee Y.C.,
Lee L.M., Luh K.T., Wu C.W.;
"Mucin mRNA expression in lung adenocarcinoma cell lines and
                                                                                                                                                           TISSUE-Breast carcinoma;
Buluwela L., Liu Q., Lugmani Y.A., Gomm J.J., Coombes R.C.;
Submitted (OCT-1992) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL.
CYTOPLASMIC (POTENTIAL).
44 X 20 AA TANDEM REPEATS.
SEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALT_INIT.
ALT_INIT.
ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; J05582; AAA60019.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA59876.1; -
AAB53150.1; -
CAA36478.1; A
CAA36477.1; A
AAB59612.1; A
CAA78972.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CAA78973.1; -
AAA35757.1; -
AAD14376.1; A
AAD14369.1; A
AAA59874.1; A
                                             SEQUENCE OF 1-89 FROM N.A.
                                                                                                                                                SEQUENCE OF 1-46 FROM N.A.
                                                                                                                             Oncology 53:118-126(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M21868; AAA58974.1;
PIR; A35175; A35175.
PIR; B35175; B35175.
PIR; S10218; S10218.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA35804.1
AAA35806.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1255
1162
1186
1255
960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24
24
1163
1187
81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        M61170; x52229; cx52228; cx
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          M31823;
S81781;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     217324;
                                                                                                                                                                                                                                                       SPLICING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
TRANSMEM
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
51;
                                                                                                               ISOPORM).
FLOTYKOGGEL -> VSIGLSFPMLP (IN SECRETED ISOPORM).
MISSING (IN SECRETED ISOFORM).
T -> A (IN REF. 11).
                                                                                                                                                                                                                                                                                                                   236 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 294
                                                                                                                                                                                                                                                                                                                                                                                             PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.9%; Score 651; DB 1; Length 1255;
27.8%; Pred. No. 9.3e-19;
ative 72; Mismatches 460; Indels 234; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPKEPAPTTPKG----TAPTTLKEPAPTTP------KKPAPKELAPTTTKGPTSTT 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SDKPAP--TTPK----ETAPTTPKEPAPTTP-----KKPAPTTPETPPPTTSEVS 820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC. . .) (POTENTIAL).
T -> TATTAPKPAT (IN ISOFORM B).
T -> TATTAPKPAT (IN ISOFORM B).
MISSING (IN ISOFORM C).
MISSING (IN ISOFORM D).
MISSING (IN POLYMORPHIC EPITHELIAL
                                                                                                                                                                                                                                                                                                                                     SAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEP-----APTTPKEPAPTTTKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                     P-KAETTTKGP--ALTTPKEP-------TPTTPKEPASTTPKEPTTI---K
                                                                                                                                                                                                                                                                                                                                                                                                                                           SAPITPKEPAPITP-----KKPAPITPKEP-----APTIPKEPTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAPTTTKEPSPTTPKEPAPTT---TKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEP--
                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----KEPAPTIKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTTPK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   357 GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             571 APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    417 GVISAPDIRPAPGSTAPPAHGVISAPDIRPAPGSTAP----PAHGVISAPDIRPAPGSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKA----AAPNTPKEPAPTTPK----EPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APTTP-----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAP
                                                                                                                                                                 T -> A (IN REF. 11).
P -> Q (IN REF. 9).
P -> Q (IN REF. 9).
S -> T (IN REF. 9).
S -> T (IN REF. 3).
A -> T (IN REF. 3).
A -> T (IN REF. 3).
                                                                                                                                                                                                                                AA; 122072
                                                                                                                                                                                                                                                                                              Conservative
  957
975
1029
1055
1133
1133
22
22
31
                                                                                                                            1087
                                                                                                                                                                            134
154
1021
1251
                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                             134
154
1021
1251
1255
                        1029
1055
1133
                                                                                                                                                     1088
                                                                                                                                                                                                                                                                                Best Local Sim
Matches 295;
CARBOHYD
CARBOHYD
                                                                                                                                                               CONFLICT
CONFLICT
CONFLICT
CONFLICT
                         CARBOHYD
                                     CARBOHYD
                                                   CARBOHYD
                                                                         VARSPLIC
                                                                                                                                                                                                                CONFLICT
                                                                                                                            VARSPLIC
                                                               VARSPLIC
                                                                                                   VARSPLIC
                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                             16
                                                                                                                                                                                                                                                                                                                                                                      295
                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                       335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             643
 Db
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

```
STTVVTTEYSTSTSTREGEITTTFVTKNIPTTYLTTIAPTP--SVTTVTNFTPTT-ITTTV 1094
                                                                                                                                                                                                                                                                                                                                                                                                                      1035 LQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQG 1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1095 IIINPMLSDETNICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIP 1154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MUCI_HUMAN STANDARD; PRT; 1255 AA.
P15941; P15942; P13931; P17626; Q14128; Q16442; Q16437; Q9Y4J2;
P15941; P15940 (Rel. 13, Created)
O1-APR-1990 (Rel. 14, Last sequence update)
20-AUG-2001 (Rel. 40, Last sequence update)
MUCIN | PRECURSOR (POLYMORPHIC EDITHELIAL MUCIN) (PEM) (PEMT)
(EPISIALIN) (TUMOR-ASSOCIATED MUCIN) (CARCINOMA-ASSOCIATED MUCIN)
REACTIVE URINARY MUCIN) (PUM) (BREAST CARCINOMA-ASSOCIATED ANTIGEN)
                                                                                                                                  767
  1142 IVVITESSTGINSA-GKITTGYTIKSVPTT-----YVITLAPSAPVIPAIN-----AVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-90368716; PubMed-2394722;
Lan M.S., Batra S.K., Qi W.-N., Metzgar R.S., Hollingsworth M.A.;
"Cloning and sequencing of a human pancreatic tumor mucin cDNA.";
J. Biol. Chem. 265:15294-15299(1990).
                                                                                                                                                                                                                                                                                                                                                                          ------TKAPKKPTSTKKPKTMPRVRKPKTPTPTPRKMTSTMPELNPTSRAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                        695 VT----SSTIESSSAPVPIPSSSTIESSSAPVPIPSSSTIESSSAPVPIPS-SSTIESSS
                                                                                                                                                       750 APVISSITESSSAPVPIPSSSITESSSAPVPIPSSSITESSSAPVPIPSSSITESSVAPV
                                                                                                                                                                                                                                                                             878 KDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITAT--TTQVTSTTTQDTTPFKI
                                                                                                                                                                                                                                                                                                                                                                                                                                              ----NSAGETTSGCSPKTVTTTVPCSTGTGEYTTEATTLVTTAVTT
                                   PEKPAPTTPEELAPT-TPEEPTPTTPEEPAP-----TTPKAAAPNTPKEPAPTTPKEPAP
                                                                                                                                712 APTTPKGTAPTTPKEPAP---TTPKEPAPT-TPKGTAPTTLKEPAPTTPKKPAPKELAPT
                                                                                                                                                                               768 TTKGPTSTTSDKPAPTTPKETAPTTPKEPAP-----TTPKKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                      810 PTPSSSSNITSSAPSSTPFSSSTESSSVPVPTPSSSTTESSSAPVSSSTTESSVAPVPTP
                                                                                                                                                                                                                                                                                                   926 KITISVITIPSITITITIVCSIGINSAGETISGCSPKIVITIVPITITISVITISITITI
                                                                                                                                                                                                                                                                                                                          936 TILKTTTLAPKVTT---TKKTITTT-EIMNKPEETAKPKDRATNSKATTPKPQKP----
                                                          PTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVT----SSTTESSSAP
                                                                                   TTPKEPAPTTPKETAPT -TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKP
                                                                                                                                                                                                                               823 T----TTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTA
                                                                                                                                                                                                                                                      870 SSSSNITSSAPSSI---PFSSTTESFSTGT-TVTPSSSKYPGSQTETSVSSTTETTIVPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SPIDTVFTRC----NCEGKT 1170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1190 TTITT--TECSAATNAAGET 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE-Pancreas;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CSTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HUMAN
                                                                                                                                                                                                                                                                                                                                                       986
                                                                                                                                                                                                                                                                                                                                                                               987
                                                                                                                                                                                                                                                                                                                                                                                                                                                   1095
                                                                                        653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οχ
                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                           g
                                                                                                                                                                                     ŏ
                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                   δ
                                         ò
                                                                                        ò
                                                                                                                                        δ
```

sequences. Differential splicing "A highly immunogenic region of a human polymorphic epithelial mucin expressed by carcinomas is made up of tandem repeats."; J. Biol. Chem. 263:12820-12823(1988). MEDILNE-9027644; PubMed-2112460; Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J., Zaretsky J., Eathen S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.; A transcribed gene, containing a variable number of tandem repeats, codes for a human epithelial tumor antigen. cDNA cloning, expression of the transfected gene and over-expression in breast cancer Lancaster C.A., Peat N., Duhig T., Wilson D.,
Taylor-Papadimitriou J., Gendler S.J.;
"Structure and expression of the human polymorphic epithelial mucin
gene: an expressed VNTR unit."; Tsarfaty I., Hareuveni M., Horev J., Zaretsky J., Weiss M., Jetsch J.M., Garnier J.M., Lathe R., Keydar I., Wreschner D.H.; "Isolation and characterization of an expressed hypervariable gene coding for a breast cancer-associated antigen."; Ligtenberg M.J.L., Vos H.L., Gennissen A.M.C., Hilkens J.; "Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino Ξ. TISSUE-Breast carcinoma:
MEDLINE=90276413; PubMed-1351132;
Wreschener D.H., Hareuveni M., Tsarfaty I., Smorodinsky N.,
Horov J., Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A., Duhig breast TISSUE=Breast carcinoma;
MEDLINE=90368715; PubMed=1697589;
Gendler S.J., Lancaster C.A., Taylor-Papadimitriou J., Duhig Gendler S.J., Lancaster C.A., Taylor Bapadimitriou J., Duhig Pendler N., Burchell J., Pemberton L., Lalani E.-N., Wilson D.; Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin.";
J. Biol. Chem. 265:15286-15293(1990). MEDLINE-88330762; PubMed=3417635; Gendler S.J., Taylor-Papadimitriou J., Duhig T., Rothbard DF3 iochem. Biophys. Res. Commun. 173:1019-1029(1990) SEQUENCE OF 1-169 FROM N.A. MEDILINE-99088473; Pubmed-2597151; Abe M. Siddiqui J., Kufe D.; "Sequence analysis of the 5' region of the human Biochem. Biophys. Res. Commun. 165:644-649(1989) MEDLINE=96183746; PubMed=8608966; Weiss M., Baruch A., Reydar I., Wreschner D.H.; may generate multiple protein forms."; Eur. J. Biochem. 189:463-473(1990). 'Human epithelial tumor antigen cDNA J. Biochem. 189:475-486(1990). carcinoma-associated antigen gene. 265:5573-5578(1990) MEDLINE=91033045; PubMed=1688329; PubMed=3417635; MEDLINE=91097524; PubMed=2268309; MEDLINE=90202794; PubMed=2318825; SEQUENCE OF 1-109 FROM N.A. PARTIAL SEQUENCE FROM N.A. carcinoma; coding for a breast-ca Gene 93:313-318(1990). SEQUENCE FROM N.A. SECUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. rissue=Thyroid; Biol. Chem. TISSUE=Breast Burchell J.; termini. tissue. Eur. J. Keydar

This SWISS-PROT entry is copyright. It is produced through a collaboration the between the Saisz Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a licensee agreement (see http://www.isb-sib.ch/announce/or send an email to licensee(abb-sib.ch). 54; CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL 1.4-LINKED ALPHA-D-GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUCING ENDS OF THE CHAINS 171 TSFDLSTGCNNYDNQCHSQTDFPGFYWNIDCDNNC------205 76 PSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSS 135 252 136 SSSSTIWKIKSSKNSAANRELOKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKV 195 196 TTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTN 255 Indels 255; Gaps 366 367 APTTEKEPAP----TTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPT 422 23 SSQDLSS-CAGRCGEGYSRDA----TCNCDYNCQHYMECCPDFKRVCTAAEVHNPTSP 75 487 P---TIPKEPAPTIKEPAPTIPKEPAPTAPKKPAPTIPKEPAPTIPKEPAPTIKEPSPT 479 534 TPKEPAPITIKSAPITIKEPAPITIKSAPITIPKEPSPITIKEPAPITIPKEPAPI-TPKKP 538 TESSSAPVISS---TIESSSAPVPIPSSSTIESSSTPVISS----TIESSSAPVPIPSSS 587 APTTPKEPAPTTPKEPAPTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTT 598 EMBL; 238061; CAA86176 1; EMBL; M16164; AAA35014.1; EMBL; M16165; AAA35014.1; EMBL; M16165; AAA35015.1; EMBL; M16165; AAA35015.1; EMBL; M18677; BAG877; BAG877; BAG877; BAG877; BAG877; ASA877; ASA878; SAB7; ASA878; KQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPT 316 TPKEPASTTPKEPTPTTIKSAPTTPKEPAPTT----TKSAP----TTPKEPAPTTTKEP N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
WW. 91C00E2DBD61AA9D CRC64; Query Match
11.1%; Score 810.5; DB 1; Length 1367;
Best Local Similarity 27.2%; Pred. No. 7.4e-25;
Matches: 332; Conservative 114; Mismatches 519; Indels 255; POTENTIAL. GLUCOAMYLASE S1/S2. SER/THR-RICH. GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REI WITH RELEASE OF BETA-D-GLUCOSE.
-! SIMILARITY: TO S.POMBE SPBC215.13.
-! SIMILARITY: SOME, TO S.POMBE SPCC285.13C. 136110 MW; 21 1367 1367 817 874 22 136 210 136 817 81 874 87 1367 AA; CHAIN DOMAIN CARBOHYD CARBOHYD SEQUENCE 206 256 338 380 423 480 a ô g δ q qq ò òγ Q οy qq g ò ò g

ŏ

ò

```
1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEQUENCE FROM N.A.
STRAIN=NCIB 10682,
MEDIATRE-9320931; Pubmed-8458832;
WEDIATRE-9320931; P., Aubert J.-P.;
"Organization of a Clostridium thermocellum gene cluster encoding the collutiosomal scaffolding protein cipA and a protein possibly involved in attachment of the cellulosome to the cellulosome to the cellulosome to the cellulosome to be surface.";
J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B) (S-LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2351
                                                                                                                                                                                                                                                                                                                                                                                          2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   981
                                                            1703 TIPSSPI-TITITPSSTITPSPPTIMITPSPTITPSPPTIMITLPPTITSSPLITIFL 1761
                                                                                                                                           1762 PPSITPPTFSPFSTTTPTTPCVPLCNWTGWLDSGKPNFHKPGGDTELLGDVCGPGWAANI 1821
                                                                                                                                                                                                                                                                                                                                                                                                                                   762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    821
                                                                                                                                                                                                                                                                       649
                                                                                                                                                                                                                                                                                                                                                     703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KPQKPTKAP-KKPTSTKKPKTMPRVRKPKTTPTPRKMISTMPELNP---TSRIAEAMLQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---GVPT-TKTPAATK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     871 PEMTITAKDKTIERDLR---TIP--ETTTAAPKMI---KETATTIEKTTESKITATTIQV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSTTTQDTTPFKITTLKTTTLAPKVTTT-KKTITTTEIMNKPEETAKPKDRATNSKATTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2236 TGTQTPTTTP--ITT--TTTVTPTPTGTQTPTTTPITTTTTTTTTPTGTQTPTTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clostridium thermocellum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
Clostridium.
                                                                                                                                                                                                                             SCRAIMYPDVPIGQLGQTVVCDVSVGLICKNEDQKPGGVIPMAFCLNYEINVQCCECVTQ
                                                                                                                                                                                                                                                                       PEKLAPTTPEKPAPTTPEELAPTTPEEPTPT-TPEEPAPTTPKAAAPNTPKEPAPTTPKE
                                                                                                                                                                                                                                                                                                            TSTTSDKPAP-TTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  THISTOLD LITTLE LANDER DE LA COMPANIA LA C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ELAPTTTKGPTSTTSDKPAP-TTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P----APTTPKEPAPT-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTKEP
                       TTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence update)
01-FEB-1996 (Rel. 33, Last annotation update)
CELL SURFACE GLYCOPROTEIN 1 PRECURSOR (OUTER LAYER PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTR-PNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTP 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2352 GTOTPITITITITIVIPIPITG---TOTP--TITPITITIVIP 2392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΑA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1664
                                                                                                                 EPAPTTP----KETAPTTP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELL SURFACE
                                                                                                                                                                                                   ---KKLTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SLP1_CLOTM
Q06852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEIN 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SLP1_CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                923
                                                                                                                                                                                                                                                                                                                                                                650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2061
                                                                                                                                                                                                      584
                                                                                                                                                                                                                                                                                                                           1882
                                                                                                                                                                                                                                                                                                                                                                                                         1942
                                                                                                                                                                                                                                                                                                                                                                                                                                                   704
                                                                                                                                                                                                                                                                                    591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                            P
                                                                                                                                                                                                   οy
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKGTAPTTLKEPAPT-TPKKPAPKELAPTTTKGPTSTTSDKPAPT---TPKETAPTTPKE 795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              289 VLAKPTP-KAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTT 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TIPKEPA--PITIKEPSPI-IPKEPAPTITKS-APTITKEPAPT---ITKSAPTIPKEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EEPAPTTPKAAAPNTPKEPAPTTPKEPA - • PTTPKEPAPT - †PKETAPTTPKGTAPTTLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EPAPTTPKKPAPKELAPTTTKEPT-STTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----PTITKEPAPT-TPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         571 APT---TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPT---TPEELAPTTPEEPTPT-TP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        348 TKSAPTTPKEPAPTTTKEPAPTTPKEPA----PTTTKEPAPTTTKSAPTTPKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KPAPTTPKEPAPTTPKEPTPTPKEPAPTTKEPAPT-TPKEPAPTAPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SDEPTPSGEPTPS--DEPTP-SDEPTPSETPEEPIPTDTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361; Indels 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1664;
                                                                                                                                                                                                                                                                                      CELL SURFACE GLYCOPROTEIN 1. 4 X 156 AA APPROXIMATE REPEATS.
ASSEMBLED INTO MONO-LAYERED CRYSTALLINE ARRAYS
                                                                                                                                                                                                                                                                                                                                                                                       O.F
               SUBCELLULAR LOCATION: CELL WALL. SIMILARITY: CONTAINS 4 S-LAYER HOMOLOGY (SLH) DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 (INCOMPLETE).
5F396695BA9FE74B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                            T-P-S-D-E-P.
GLY/PRO/SER/THR-RICH.
SLH 1 (INCOMPLETE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13.0%; Score 950; DB 1; 31.0%; Pred. No. 3.8e-30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match
13.0%; Score 550, DB
Best Local Similarity 31.0%; Pred. No. 3.8e-
Macches 312; Conservative 104; Mismatches
                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                           EMBL; X67506; CAA47841.1; ...
Interpro; IPR001119; SLH.
Pfam; PF00395; SLH; 3.
PROSITE; PS01072; SLH_DOMAIN; 2.
PROSITE; PS01072; SLH_DOMAIN; 2.
SIGNAL
                                                                                                                                                                                                                                                                 .qnal; Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                          SLH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          178194 MW;
                                                                                                                                                                                                                                                                                                              763
191
363
565
763
                                                                                                                                                                                                                                                                                                                                                                                                                                            1494
1565
1625
1646
                                                                                                                                                                                                                                                                                                1664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                            1453
1495
1566
1626
1664
                                                                                                                                                                                                                                                                                                                                                              409
                                                                                                                                                                                                                                                                                                29
36
36
207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    624
                                                                                                                                                                                                                                                                                                    CHAIN
```

aa6.rsp

2

```
CC the State of the Copy of th
```

	52:											
	Sabs	, ,	115	150	181	213 1363	241	301	348	398 1582	451	80 07
TIAL). TIAL). TIAL). TIAL). TIAL). TIAL). TIAL). TIAL). TIAL).			ESEE ::: TADKCG	KSSKNS	i. Š	SPK	SLTVN:	(AETTT 	APTTT PPTTT	TTPKEPA	APTAPKK PPTTTPS	AP TPSPT
(POTENTIAL)	th 517 els 3	PDFKRVCTAAEVHNPT 	CTKKVI :	WKIKSS : CYWEI	VV	-NKVSTSPK GQKVQCDVS	DTSKE1 : TTTLP	AKPTPKA! : -TPSPPI	 PASTITLPPIT	T PTT	AP TTTLPP	PTTTKS : PTTTMT
CRC (Length	CPDFKR	-PPNKKTKKVIESEE :: :::: RPIYEEDLKKCVTADK	TODGAE	PTPKPE TTPTSS	STTQH- SLEQHG	SLPPNSDTSKETSLT 	APTSKVL	-TTPKEP- MTTPITPP	EPAPTTTKSAP : SPPPTTTTPP	(EP TPPTS	TKEPA TTPSP
(GLCNAC. (GL	DB 1; e-38; 524	C RTPDLC	KRSPK PRCPKDF	QESSSSSSSSSSTI	KLKVKDNKKNRTKKKPTPRPPVV 	VTTPDTSTTQH-NKVSTSPK 	INPRP PTTTP	STDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTT :		PTTTKEPAPTTTKSAP : PTTTPSPPPTTTTTPP	EPAPTTPKEPAPTAPK SPPTTTPITPTSTTTLPPTTP	SPTTPKEPAPTTKSAPTTTKEPAPTTTKSAP
	1171; No. 3.1	3ACVFW	EGCYP	CRPEE	KDNKK TPTSF	IECRS	ITTAK TTPS	SIEKT	TPTTI PPTTI	APTTTK STTTLPPTTTP	APTTKEF PTTTPSF	APTTT
LLINK	Score 117 Pred. No. 9; Mismat	EC ECTKE(ISVSYI	SSSSTI::	KLKVKDN TTITLPTTPT	CGAPED	*PMDKC	SAKETQ : !TTP	PTTPKEPASTTPKEPTPTIKSAP 	r T	PKEPA TTPPP	TTPKEI TTTTT
DE MANAGEMENT OF THE PROPERTY		NCQHYM : VASYAQ	r	SSSSSS	STLTTFT	PFDGVCGA	RVNCC	KEKTT	KEPAS: : PSPPT	EPAPTTPKEP SPPTTTPITPPT	EPTPTI PPTTT	TKEPSF PSPPI
4339 4351 4352 4423 4423 4423 4420 4616 4616 4787 4787 4787 4787 4787 4787 4787 478	16.1% 29.2% ative	SRDATCNCDYNCQHYMECCPDF 	PPSGASQTIKSTTKRSPKPPNKKKTKKVIESBE- : : :	SENQESSSSS : : VPTEETCKSC	TTRPS	EAGSGLDNGDFK- : : HPSSGSDDGDREP	CKNEDQFGNGPFGLCYDYKIRVNCCWPMDKCITTPSP	NKQTSTDGKEKTTSAKE : PPTTTPSPPITTTTP	SP	PKEPAPTTKEPAPTTPKEP 	PAPTTREPAPTTPREPTPREPAPTREPAPTTPEP 	PAPTIPKEPAPTITKEPAPTITKEPSPTIPKEPAPTITKSAPTITKEPAPTITKSAP
Ž.	larity Conserva	YSRDA: GDCI	PSGASC : : SNRSFE	SVSEN : PPGASVPT	FNICSI	35	GPFGL	ETTTTNK -TTTTPP	TIP	PAPITI PPPIT	TPKEP.	PTTPK PTTTP
4339 4351 43733 43733 44522 44528 4452 4452 4461 4688 4688 4697 6501 6501 1351 11412 11504 4192	1mi ,	RCGEGY:	SSKKAPPP EWHYEPCGI	EEH- : DTHY	. =	DEZ DWINEDHI	EDQFGN	KETTVETKETTTTNKOT 	<u>a — a </u>	SAP-TTPKEPAPTTK 	PKKPAPT PSPPTIT	PKEPA TTPP
CARBOHYD CAR	atch cal 39	9 SCAGRO : 9 SCSCI	5 PPSS 9 ECEW	6IT. 9 CYVE	AA GP	LWS	•		KGPALTT TPPPTTT	× σ	PTTPK TTTPS	PAPITI PPPIT
	uery est L atche	1079	75	116	151	182	214	242	302	349	399 1583	452
747 747 747 747 747 747 747 747 747 747	ÖÄÄ	Qy Db	QY	Qy Db	Qy	Oy Op	Oy Dp	Qy Db	oy Op	o da	Oy Op	O _Y
												

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

April 26, 2002, 16:32:37; Search time 62.75 Seconds (without alignments) 798.154 Million cell updates/sec Run on:

AA6 7294 1 MAWKILPIYLLLLSVFVIQ.....ARAITTRSGQTLSKVWYNCP 1366 BLOSUM62 Gapop 10.0 , Gapext 0.5 Title: Perfect score: Sequence: Scoring table:

100059 seqs, 36664827 residues Searched: 100059 Total number of hits satisfying chosen parameters:

 $\begin{array}{c} {\tt Minimum\ DB\ seq\ length:\ 0} \\ {\tt Maximum\ DB\ seq\ length:\ 200000000} \end{array}$

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:* Database :

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description			saccharomy		mytil	Q9y493 homo sapien			P23253 trypanosoma		_			рошо		_		_	mod1u		_	-		-					schizos	mytil	рошо	3428 homo	homo	
	ID	MUC2_HUMAN	SLP1_CLOTM	AMYH_YEAST	MUC1_HUMAN	FP1_MYTED	ZAN_HUMAN	EXTN_TOBAC	NFH_MOUSE	TCNA_TRYCR	CPN_DROME	FP1_MYTCO	NFH RAT	MUC1_XENLA	RPB1_HUMAN	NFH_HUMAN	RPB1_MOUSE	RPB1_CRIGR	BAT2_HUMAN	SSP2_PLAYO	EXTN_MAIZE	ZAN_MOUSE	HWP1_CANAL	VRP1_YEAST	VGLX_HSVEB	YJ9P_YEAST	N121_HUMAN	VG50_HSVI1	MAPA_RAT	YAV1_SCHPO	FP1_MYTGA	T2D3 HUMAN	TCOF_HUMAN	KI67_HUMAN	
	DB	1	Н	-	Н	Н		Н	Н	Н	,_	Н	-	,-I	Н	Н	Н	Н	Н	Н	Н	Н	Η	Н	7	Н	Η,		-	-	-	7	٦	Н	
	Length		1664	1367	1255	875	2700	620	1087	1162	865	872	831	662	1970	1020	1970	467	2142	826	267	5376	634	817	797	1161	1229	670	2774	1794	751	1083	1411	3256	
	Match	16.1	ϵ	11.1					7.6	7.3	7.3	7.1					9.9					•		6.4	6.2		6.2	6.1	6.1	- C	9.1	0.9	0.9	6.0	
	Score	17	950	810.5	651	559	755.5	, , , ,	551	530.5	530	518.5	517.5	5.005	503.5	502	498.5	493.5	-	488.5	475.5	-	468.5	467	454	454	453.5	48			٧ ۲	0		35.	
Result	NO.	-	17	m	4	٠ ١٠	ı ve	2	· oc	σ	, [3 -	17	7 -	14	· ·	16	17	18	19	20	21	22	23	24	25	26	22	ď	000	(e	3.5	3.5	3.6	

P17437 xenopus lae P10220 herpes simp Q99618 h nuclear r Q09550 caenorhabdi Q9umn6 homo sapien P5425 homo sapien P27546 mus musculu P27546 mus musculu P28955 equine herp P54258 rattus norv Q2893 sus scrofa P03200 epstein-bar	
XP2_XENLA TEGU_HSY11 VGU3_CAEEL YRX2_HUWAN VGU3_CAEEL TEX2_HUWAN MAP4_MOUSE SGS3_DROME TEGU_HSVEB TEGU_HSVEB ZAN_PIG VGP3_EBV	
нанананана	
439 3164 2517 12517 2715 1185 1125 307 3421 1183 2476	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	
432.5 423.5 427.5 427.5 425.5 424.4 423.5 419.5	
44444444444444444444444444444444444444	

ALIGNMENTS

. ; * . . • •

```
A; Residues: 1-489 <ARS>
A; Cross-references: EMBL.X91836; NID:g1015936; PIDN:CAA62943.1; PID:g1015937
A; Cross-references: EMBL.X91836; NID:g1015936; PIDN:CAA62943.1; PID:g1015937
A; Experimental source: sub_species Red calcoon
R; Arsenjevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
Submitted to the EMBL Data Library, April 1995
A; Description: A class of root-hair specific extensins involved in rhizobium/legume inte
A; Reference number: 554155
A; Accession: 554155
A; Accession: 554155
A; Accession: 5480
A; Cross-references: EMBL:X86030; NID:g791149; PID:g791150
C; Genetics: A; Genetics: A; Genetics: A; Genetics: A; Genetics: G; Cyperfamily: hydroxyproline
C; Superfamily: hydroxyproline-rich glycoprotein
C; Keywords: glycoprotein; hydroxyproline
F; 1-23/Domain: signal sequence #status predicted <MID: F; 24-489/Product: extensin class 1 #status predicted <MID: A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    637 NTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELA 696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      697 PITIKEPISITSDKPAPTIPKGTAPTIPKEPAPITPKEPAPTIPKGTAPTILKEPAPTIP 756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457 PKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSP 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              517 ITTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTKKPAPTAPKEPAPTTPK 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              577 ETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAP 636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   248 PPPPYYKKSPPPPSPSPSPPPSPSPSPPPYYKKSPPPP 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PSPSPPPPYYYKSPPPPSPSPSPPPYYYKSPPPPSPSPSPPPPSPSPSPPPYYY- 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          400 TTPKKPAPTTPKEPAPTTPKEPTTPKEPAPTTKEPAPTTP---KEPAPTAPKKPAPTT 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142 YKSPPPPSPSPPPPYYKSPPPPSPSPPPYXYKSPP------PPSPSPPPP 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            340 PKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAP 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             39 PKQTPPYYYNAPPYYYKSPPP-----PSPSPP-----PPPYVHKYPPYYKKSPPP 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPP 814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match

8.7%; Score 633; DB 2; Length 489;
Best Local Similarity 31.2%; Pred. No. 2.8e-20;
Matches 149; Conservative 33; Mismatches 252; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŻ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

Search completed: April 26, 2002, 16:32:20 Job time: 735 sec

230 225 225 288 280 280 344 340	Db 391 SDSSSNEETTTKPAAKTTPAKSA-ATPTSKTPTNGK-ATPTSKTPAKCTPAKGTPAKKDS 448 Qy 451 KPAPTTPKEPAPTTTKEPSPTTFKEPAPTTTKSAPTTTKSAPTTTSAPTT 510 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Db 644 PAVKTLPPKKAESSSDSSKTKTKPARPPARSATPVNTKAPAQNKASKASSDSS 703 Qy 743 TAPTILKEPAPTTEKPAPELEDTTTKGPTSTTSDRPAPTTTREPAPTTPK 802	RESULT. 15 Til622 extensin class I precursor - cowpea Cispecies: Vigna unguiculata (cowpea) Ristenijevic-maksimovic, I.; Broughton, W.J.; Krause, A. Mol. Plant Microbe Interact. 10, 95-101, 1997 A;Title: Rhizobla modulate root-hair-specific expression of extensin genes. A;Reference number: 217301; MUID:97155574 A;Accession: Til622 A;Status: preliminary; translated from GB/EMBL/DDBJ
2417 647 2473 682 2533 716 2593 765	Qy 824 TTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMT 874 bb 2713 STSSEPETHAPATYSEASSTTLEENSSTSSPTSSEASVKLSSLFPESI 2762 Qy 875 TTAKDKTTERDLRTTAAPKWTKETATTTERTESKITATT	RESULT 14 ISSULT 14 INCLEDIAT Phosphoprotein - African clawed frog C; Species: Xenopus laevis (African clawed frog) C; Species: Xenopus laevis (African clawed frog) C; Species: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999 C; Accession: 151618, 557757 C; Accession: 151618, 557757 J. Cell Sci. 108, 3339-3347, 1995 A; Title: Identification and CDNA cloning of a Xenopus nucleolar phosphoprotein, xNoppl86 A; Reference number: 151618, MUID: 96019267 A; Accession: 151618 A; Residues: preliminary; translated from GB/EMBL/DDBJ A; Ctatus: preliminary; translated from GB/EMBL/DDBJ A; Ctors references: EMBL: X88927; NID: 9895920; PIDN: CAA61881; PID: 9895921 A; Genetics: C; Genetics: A; Genetics: C; Superfamily: nucleolus-cytoplasm shuttle phosphoprotein C; Superfamily: phosphoprotein	Ouery Match Best Local Similarity 24.2%; Pred. No. 26-20; Matches 246; Conservative 148; Mismatches 424; Indels 199; Gaps 40; Matches 246; Conservative 148; Mismatches 424; Indels 199; Gaps 40; Qy 74 SPPSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEE 119 L

```
0 SDISKETSLIVNKETIVETKETITINKQISIDGKEKITSAKETQSIEKTSAKDLAPIS-- 287
                                                                                       8 --KVLAKPTPKAETTTKGPALTTPKEPTPTTPKE--PASTTPKEPTFTIKSAPTTPKEP 343
                                                                                                                                                                        4 APTT------TKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKE-PAPTTTKSAP 391
                                                                                                                                                                                                                                                                      0 DSSSEDEKKSSVKLGVKAAP--KKAPA----APDAKSTPVAAAKKSAPAKKASSSSD 390
                                                                                                                                                                                                                                                                                                               2 ТТРКЕРАРТТРККРАРТТРКЕРАРТТРКЕРТРТТРКЕРАРТТКЕРА-РТТРКЕРАРТАРК 450
                                                                                                                                                                                                                                                                                                                                                                                                    1 KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTT 510
                                                                                                                                                                                                                                                                                                                                                                                                                                  class 1 precursor - cowpea
f. 'Yana unguiculata (cowpea)
6-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
on: T11622; S54155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTFEEPAPT 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              О ТРКАААРИТРКЕРАРІТРКЕРАРІТРКЕРАРІТРКЕТАРІТРКСТАРІТІКЕРАРІТРКК 689
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SAKPAVKIT---PGKATSK---PVVASKPVP-AKKASSSSDSDSSEEETIKITKPLTKLS 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAPKELAPTTTKEPTSTTSD-----KPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKG 742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEBEGKSKQP-TGKSPAAKATAP-PKKNPVAVNKDKPSSSSSSSSSGDDEKQ----KPK 757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KPAPTIPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPT 862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SDEDVSKAKKTNTAVSKSPV----TTPKAVPAKKESSSESSDSEDEKGGGKNTSTTKI 843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TSTTTQDTTPFKITTLKTTTLAPKVTTTK--KTITTTEIMNKPEETA-KPKDRAINS--- 976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 TAPITLKEPAPITPKKPAPKELAPITIKGPTSTISDKPAPITPKETAPITPKEPAPTIPK 802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TKIPAAIKPEMITIAKDKITERDLRITPETITAAPKMIKETAITIEKITESKITATITQV 922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ANST------PKAAAAECSEESSSSEDEGKANGTSGKRKRESTGNAEC 885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTP-RKMTSTMPELNPTSRIAE 1032
                                                                                                                                                                                                                                                                                                                                            PKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEP-APTTPKEPAPTTKKPAPTAPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S
```

yo 4	478		ഗ
a A		APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTKKPA- 563	
qq	: 2346 TT	: : : : :	o
οy	564	588	ć
QQ	2400 RS	ETVHTSTVLTTTAT 245	on.
Qy	589	628	,
QQ	2460 TT	251	o.
Qγ		PKEPAPTTPKETAPTTPKGTAPTTL 679	c
qq	2520 TV	SMATPSSSTQTSGTPPSLITTATTITATGSTT = 23/	œ.
οy	089	KEPAPTIPKKPAPKELAPT-TIKEPISTISDKPAPTIPKGTAPTI /23	
qq	2579 -NP	SSTPGTTPIPPELTTTATTPAATSSTVTPSSALGITHIPPVFNITALINGKSLSFSS	
δ δ		PTIVKKY 739	7.0
2	r E	:	
δλ	760 API -	KELAPTTTKGPTSTTSDKFAFTTFKETAFTT-FREF 	
qq	2698 SP	GTTTPGHTTATSRTTATATPSKTRTSTLLPSQPTSAPITTVVTTGCEPQCAWSE	5 .
οy	197	796	10
QQ	2754 WL	WLDYSYPMPGPSGGDFDTYSNIRAAGGAVCEQPLGLECRATAQPGVPLGELGQVVECSLD 281:	13
QY	797	APTTPKKPAPTTPETPPTTSEVSTP 822	0
QΩ	2814 FG	r 287	7.3
οy	823 TT	DESTP	S
qq	2874 AT		30
Οy	867 AA	VIKTAAPKMIKETAT 905	S.
op	2931 TA		06
ò	TI 906	PEKTFKITTLKTTTLAP 945	ZI
7 A			50
Qy	946 KY	66	2
qq	3051 GS		10
δy	993 PJ	PISIKKPKIMPRVRKPKTIPIPRKMIST 102	20
QQ	3111 A	: : : :	70
Qy	1021 M	MPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR 107	72
Op	3171 TI	TPETVHISTVLITITATITGATGSVAIPSSTPGTAHTIKVPITITIGFTATPSSSPGTALT 3230	30
٥y	1073 PI	RVPNQGIIINPMLSDETNICNGKPVDGLTTLRNGTLVAFRGHYF 11	32
qq	3231 P		180
Qy	1133 W	WMLSPFSPARRITEVWGIPSPIDTVFTRCNCEGKT 1170	
ç	3281 -	: : : : : : : : : :	

															-		
50;																	
Gaps	115	175			w 1	368		477	7 2	593	644	98 4	758	809	860	916	970
962;	ESE : LPKEE	ITEBHSVSENQESSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKRPT 	SLPP	EVEEFI	<u>م</u> ي	EPTPTTP-KEPASTTPKEPTPTTKSAPTTPKEP-APTTTKSAPTTPKEPAPTTKEPAP 	KEPTPT : EVPKAA	TTKEPS : KKVPAA	PTTPKK : VAVPEK	-PTTPEK KNPVPEK	AP VPEKKV	KELAPT	PTTPKK : APPPKE	-APTTP (VELTP	KEPGV : KKPAV	ESKIT	
Length 79	PDEKRVCTAAEVHNPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVT : : :	KDNKKNE : PEVIKKA	PKPFVVDEAGSGLDNGDFKVTTP-DTSTTQHNKVSTSPKITTAKPINPRPSLPP 	-NSDTSKETSLIVNKETTVETKETTTINKOTSTDGKE	TTTKGPALTTPK : KGPEISEKIIPPKK	KEPAPTT	SAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPT 	TPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPS	PTTPKEPAPTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPK 	KEPAPTTEKEPAPTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEK 	LAPTTPEKPAPTTPEELAPTTPEEPTP-TTPEEPAPTTPKAAAPNTPKEPAP 	-APTTREPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT	TTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKK	PAPKELAPTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTP 	ETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGV	PTTKTPAATKPEMTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKIT 	FKITTLKTTTLAPRVTTTKKTITTTEIMNKPEETAKPK
	PKPPNKI : PEVPKK	ZKKLKVI : I VAPPRVI	ITTAKP. KVPEKKN	IEVEEYE	AETTTKG KG	SAPTTPK EVLPPK	APTTPKE : PKKPEV	PAPTTP . RVPEVP	KEPAPT 	PTTPKK : PAAPPQ	PKAAAP PPAKVP	EPAPTT SPPPEV	KGTAPT' : EKKVP-	KEPAPT : KKKVPE	SAEPTPI -	TKETAT PAPEPKE	EMNK
; DB .9e-20 es 48	STTKRS:	AANREL(EEVEE'	VSTSPK: VPEVPKI	rdgke :: PEEEEF1	AKPTPK! APP	APTTTKE PPTKVE	()	APTTPKE	PSPTTT PEVPPT	TPKETA : I :VVEEPE	EPAPTT : KPEA	APTTLK : VPERAE	PAPTTP PKKIIP	PTTP 	ESTPEL: :: EEEPKVI	TTAAPKI KPAIPLI	LITTTE: : : PVTVPV
ore 661.5; DB 2; ed. No. 2.9e-20; Mismatches 480;	ASQTIK	KSSKNS PEEPPS	TTQHNK' EAPPAK'	-NSDISKETSLIVNKETIVETKETITITNKOISTDGKE : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	KTTSAKETGSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALT 	EPTPTTP-KEPASTTPKEPTPTIKSAPTTPKEP-APTTKSAPTTPKEPAP 	TTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKK 	TAPKKPA: 	PITPKEPAPITIKSAPITIKEPAPITIKSAPITPKEPSPITIKEP 	PAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLT 	P-TTPEE	TTPKGT AVP	PTTPKE :{ PAVPEE	A	IHKSPD: : : LRKRPE	RTTPET TKVPEI	ATTTQVTSTTTQDTTPFKITTLKTTLAPKVTTTKKTITTTEIMNK : : : : : : : : : :
	PTSPPSSKKAPPPSG :: VIPKKEEAPP	SSTIWKI :!	TP-DTS VPVQKK	ETTT ELEVVE	SAKDLAI ::: EAEEVE	IKSAPT: :: VEEKVR	TPKEP! : (VPEAP!	KEPAPI : :	TTTKS?	(PAPTAF PEEVAL	TPEEPT : VPEKKV	PKETAP -LEEKP	TPKEPA : EPEVPP	TTPKET APPKED	TT : VVLKSV	TTERDL : RKIPEP	TTLAPK : : PPVEPE
9.1%; Sc 24.3%; Pr tive 121;	PPSSKK :: IPKKEE	SSSSSS	GDFKVT	IVETK- :: VVKERL	RVIEVE	SPTPTT:	TTKSAP1 PPF	PPEVFER	TKEPAE	PTTTKK :: PEVFEE	EELAPT EVPKEV	EPAPTT : KPV	KGTAPT : EE	SDKPAP PPPPPP	ETPPPTTSEVSTPTTTKEP 	TTAKDK : TVTKRE	ITTLKT: : : VKTIKP!
Similarity 246; Conservative	VHNPTS VV	ITEEHSVSENQESSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	GLDN : KVSKKI	TVNKET: : AFRQEV	KTTSAKETQSIEKT : : : KLEQHRVEEEHRVEKVHRV	ASTTPKI : : PAKVPEV	CEPAPTI CPEAP	APTTP- : KPESPF	KSAPTT K-VPEP	TPKEPA : { KPESPP	PAPTTP PPVKVP	APTTPK VPEVPK	PAPTTP PVPVAE	GPTSTT PKLKPR	-TPTTT LLPERK	EMT : EEVEVP	DTTPFK:
milari Cons	PDFKRVCTAAEVH : : : PEIKKKVTEKKV-	SVSENQ : VTEEPE	/DEAGS(::: /EAPPAI	SKETSL: :: : SEKVSV	SAKETQS : :VEEEHI	P-KEP	TTPKEPAPTTTKEPAPT 	TPKEPAPTTKEPAPTTP : : VPEKKVPEAIPPKPESP!	EPAPTTI : KPEVTPV	KEPAPT 1 -EAIPP	PEK AKKPEL	EVPPAK	STTSDKP.	APTTK IEK	TSEVS- : EKKVRK	AATKP-	STTTO
1tch 26	PDFKR : : PEIKK	ITEEH: : VLPVEY	PKPPV1	-NSDTS : GRTVLE	KTTS : KLEOHF	EPTPTT PPTKVV	TTPKEP -EKKVP	TPKEPA :: VPEKKV	PITPKE : PPKK	PAPTTP	LAPTT 	TTPKEPAPTTPKEPAP : : AVPKKPEVPPAKVPEVPKRPV-	TTKEPTSTTSDKPAPTTPKGTA : : : : EEEIAPEEEKPVPVAEE	PAPKEL ::: PEPEKV	ETPPPT: LKVPGG1	PITKIPA PEPI	ATTTOV
Query Ma Best Loc Matches	58 6853	116 6897	176 6948	229	265	311	369	425	478	538 7336	594	645	699 7	759	810 I	861 E	917 #
OME	o do	og G	Oy Op	Qy Op	oy ob	9 9 P	oy ob	oy g	oy op	oy op	\$ a	ço qa	o o	Op Op	oy op	Oy Dp	Oy Dp

Cispecies: Homo sapiens (man)
Cispecies: 1-Jan-2000
Cispecies: Homo sapiens (man)
Cispecies: Homo sapiens
Cispecies: Homo sapiens
Cispecies: T45025
Cispecies: Mise-3178, 1997
Cispecies: Mise-3178
Cispecies: 64; 7758 GPIKGVPKKTPSPIEAERRKLRPGSGGEKPPDEAPFTYQLKAVPLKFVKEIKDIILTESE 7817 971 DRAINSKAITPKPQKPIKAPKKPIS--IKKPKIMPRVRKPKITPTP-----RKMISIMPE 1023 1693 RTATTLPVLTSTATKSTATSFTPIPSSTLGTTGTSQNRPPHPMATMSTIHPSSTPETTHT 1752 1326 LTTTATTTRIGSVATPSSTPGTAHTTRVPTTTTTGFTATPSSSPGTALTPPVWISTTTP 1985 62 RVCTAAEVHNPTSPPSSKKA--PPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEE 119 120 HSV----SENQESSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRT--KK 172 2226 LDYSYPMPGPSGGDFDTYSNIRAAGGAVCEQPLGLECRAQAQPGVPLRELGQVVECSLDF 2285 221 NPRPSLPPNSDISKETSLIVNKETIVET-----KEITTINKQISTDGKEKTISAKETQSI 275 276 EKTSAKDLAPTSKVLAKPTPKAETTTK---------GPALTTP----KEPTP 314 Gaps KEPAPTTKEPAPT--TPKEPAPTAPKKPAPTTPKEPAPTT-PKEPAPTTTKEPS----- 477 173 KPTPKPPVVDEAGS-----GLDNGDFKV----TTPDTSTTQHNKVSTSPKITTAKPI 344 APTITKSAP--TIPKEPAPTIT-KEPAPT----TPKEPAPTITKEPAPTITKSA-----391 PTTP----PKEPAPTTPKKPAPTT----PKEPAPTTPKEPTT Query Match 8.9%; Score 649; DB 2; Length 3570; Best Local Similarity 22.5%; Pred. No. 4.38-20; Matches 373; Conservative 111; Mismatches 584; Indels 590; mucin MUC5B, tracheobronchial [imported] - human (fragment) 1024 LNPTSRIAEAMLQTTT 1039 :| | | :: :| 7818 FVGSSAIFECLVSPST 7833 478 427 qq δ Ω ö ö q ŏ g δλ QQ òγ Dp δ g ò g g à D å ò

```
A;Gene: GDB:TTN
A;Cross-references: GDB:127867; OMIM:188840
A;Map position: 2q31-2q31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      889
                                                                                                                                                                                                                                                                                                                                                              545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      937
                                                                                                                                                                                                                                                                                                                                                                                                           969
                                                                                                                                                                                                                                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                                              649
                                                                                                                                                                                                                   609
                         504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óγ
                                                                                                                                                              Ω
                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                 οχ
                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Gene: GUB:MUCLI; FUM
A;Gene: GUB:MUCLI; FUM
A;Gross-references: GDB:120705; OMIM:158340
A;Gross-references: GDB:120705; OMIM:158340
A;Torsor: 20/1; 62/3; 1165/3; 1184/2;
A;Introns: 20/1; 62/3; 1165/3; 1184/2;
B;1-1344/Product: mucin l precursor, splice form A *status predicted <PREA>
F;1-12/2/Domain: signal sequence #link PREB #status predicted <SIGB>
F;1-19,29-32/Domain: signal sequence #link PREB #status predicted <SIGB>
F;1-19,29-124/Product: mucin l precursor, splice form B *status predicted <PREB>
F;1-19,29-1134/Product: mucin l precursor, epithelial tumor antigen splice form F;138-1017/Region: Mucin l carboxyl-terminal non-repetitive
F;138-1017/Region: mucin l carboxyl-terminal non-repetitive
F;1245-127/Domain: transmembrane #status predicted <FRM>F;1046-1118,1144,1222/Palinding site: carbobydrate (ASn) (covalent) #status predicted
F;1213/Binding site: phosphate (Tyr) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                             A.Title: Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins cytokine A.Feference number: 551026; MUD: 95080414
A.Contents: annotation
A.Kontents: annotation and tyrosine residues in the carboxyl-terminal non-repetitive region are cycomment: This protein is length polymorphic. Individuals may have between 21 and 125 c partial repeats. The repeat shown is defined by Smal nuclease sites.
C.Comment: Scrine and threonine residues in the tandem repeat domain are extensively gly cycomment: For an alternative splice form without a tandem repeat domain, see PIR: 540146 A.Gene: GDB: WUC1; PUM.
                                                                                                                                     A; Residues: 1-142,'0', 144-162,'0', 164-168 <ABE>
A; Cross-references: EMBL:M1823; NID:9181542; PIDN:AAA35757.1; PID:9181543
A; Cross-raferences: EMBL:M1823; NID:9181542; PIDN:AAA35757.1; PID:9181543
B; Masuzawa, Y.; Miyauchi, T.; Hamanoue, M.; Ando, S.; Yoshida, J.; Takao, S.; Shimazu, H. S. Blochem. 112, 609-615, 1992
A; Title: A novel core protein as well as polymorphic epithelial mucin carry peanut agglu A; Reference number: JX0235; MUID:93123189
                                                  human DF3 breast carcinoma-associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56;
                                                                                                                                                                                                                                                                                                                                                           A:Experimental source: gastric carcinoma cell
R:Zrihan-Licht, S.; Baruch, A.; Elroy-Stein, O.; Keydar, I.; Wreschner, D.H. FEBS Lett. 356, 130-136, 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSA 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---KKPAPTTPKEP---- 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            307
                                                                                                                                                                                                                                                                                                                      A;Molecule type: mRNA
A;Residues: 998-1011,'ES',1014-1017;1018-1032,'T',1034-1037;1038-1057 <MAS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LTVNKETTVETKETTTTNK--QTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---KE 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPAPTTPKEPAPTTPK - - - - EPAPTTTKEPSPTTPKEPAPTT - - - TKSAPTTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LIVLIATTAPRPATVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---KEPAPTTKEP-----APTTPKEPAPTAPK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches 472; Indels 256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 1344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----KSAPTTPKEPAPTTTKEP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-KAETTTKGP--ALTTPKEP--------TPTTPKEPASTTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9.2%; Score 672; DB 1;
27.8%; Pred. No. 1.7e-21;
iive 78; Mismatches 472;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP-----
A,Title: Sequence analysis of the 5' region of the A,Reference number: A36735; MUD:90088473
A,Rocession: A36735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PTPTT----IKSAPTTPKEPAPTT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 27.8 nes 310; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --APTTPKEPTPTTP-
                                                                                                                                   A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                         A; Accession: PX0066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

```
Rilabeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
Aritile: Titins: giant proteins in charge of muscle ultrastructure and elasticity. A; Reference number: A57430; MUID:96026330
A; Accession: 138346
A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g1017427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         elastic titin - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1004 AHGVISAPDIRPAPGSTAPPAH-----GVISAPDNRPALGSTA----PPVHNVIS 1049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STKKPKTMPRVRKPKTTPTPRKMTSTMPELNP----TSRIAEAMLQT-TTRP---NQTPN 1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      947 VISAPDIRP---APGSTAPPAHGVISAPDIRPAPGSTAPPAHGVISAPDIRPAPGSTAPP 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1047 SKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSDETN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----TKPEMTTTAKDK---TTERD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    886 LRTTPETT-----TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTP--FKIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TLKTTTLAPKVTTTKKTTTTT - - EIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAP--GSTAPPAHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           833 TRPAPGSTAPPAHGVTSAP----DTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EPAPTTP-----KKPAPKELAPTTTKGPTSTTSDKPAP--TTPK----ETAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        656 APGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTR
                                                                                                                                                                                                                                                                                  HGVTSAPDTRPAPGSTAP----PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTPKEPAPTTPKG----TAPTTLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AP-PAHGVTSAPDTRPAPGSTAPPAHGVTS----APDTRPAPGSTAPPAHGVTSAPDTRP
--APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK-
                                                                                                                                                                                 HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                      -KKPAPKEL
                                                   ---EPAPTITKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPE-
                                                                                                                                                                                                                                                 ---ELAPTTPEEPTPTTP-----EEPAP--TTPKA----AAPNTPKEPAPTTPK-
                                                                                                                                                                                                                                                                                                                                                                         ---EPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --KKPAPTTPETPPPTTSEVSTPTTKEPTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1107 ICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPS 1142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ASGSASGSASTLVHNGTSARATTTPASKSTPFSIPS 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: mRNA
A; Residues: 1-7962 <RES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C; Accession: I38346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEPAPTTP-
```

Page 8

oy op	2 AWKTLPIYLLLLSVFVIQQVSSQDLSSCAGRCGEGYSRDATCNCDYNC 50 : : : : : : :	
δ	PSGASQTIKSTTKRSPKPPN 1	
qa	IGPSGDSNPTGSSWFQEIEAT 98	
òá	4 KKKTKKVIESPEITEEHSVSENQESSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK	
3		
දු දු	NKVSTSPKITTAKPINPR	
ò		
7 음	172TTELLISPETLVSTRSTVSTSAKETGSIEKTSAKET 283 172TTELLISPETLVSTDSSTST-EOTSPDNITTERSPERENTERS	
δ	, r	
QQ	2 5	
Qy	40	
qq	259 AESTTKKSTTKAPATTEEPTPTTEEVTTTEAETSTTTSSETSTEKPTTPL 309	
ογ		
qq	310'IDNKIAGPATGKPETTHFPVTGTTPNFDTATETPFVAKSEDKMTLSKTAATETTQ 364	
ò	459 EPAPTTPKEPAPTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPT 502	
q	365 QTTEVT-DGPEKETTKNVSIEIPITTVPLVETTSTASKESDGFHTTLKLKVTTADSD 423	
ò	503 TTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAP 540	
đ	424 STESATTVKPFNEETTTKSHVVPKPTKKGTVKVTPKLELSFDEPTEIT-KAPHPGKLLEK 482	
ολ	541TTPKEPAPTTPKEPAPTTTKKPAPTA- 566	
q	483 KTYHFVLSDNFARYSEAKENDDYNHLDYNHYREAKEPTTTEESSTTEEVTTTEEPANTGN 542	
ŏ	EELAPTTPEEPT 61	
qq		
δ		
QQ	599 SIEKTSTIKKASITEEPUTITDEPUTITESSITGKATIPELSTISEETIT 647	
ģ	PAPKELAPTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEP	
д	69	
δ	Δ.	
g		
δ	۵	
qq		
ογ	815 TISE-VSTPTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEM 873	
g G	œ	
ò	TTTAKDKTTERDLR	
q		
ογ	896 -APKMIKETATITEKITESKITATITQ-VTSTTIQDTIPFKITILKITILAPKVT 948	

q	
, QY	PEETAKPKDRATNSKATTPKPQ
qо	968 PVDETTINTLELLSKINNTQISQPKPTDISKTDALSSLISGLIGSFTKAP 1017
Qy	033
RESU	LT 10
mucin 1	ucin 1 precursor, repetitive splice form A [validated] - human
n; AI ncre	DF3 antigen; cor EM)
C; Sp	antigen splice fo
C; Ac R; Li	C. Accession: A35175; B35175; A35886; A35887; S10572; \$40293; A36735; PX0006; S10218; R; Ligtenberg, M.J.L.; Vos, H.L.; Gentlasen, A.M.C.: Hilkens
A;Ti	n, is denerated by a nolumorphic con
A; Re A; Ac	
A; Mo A; Re	lecule type: mRNA sidues: 1-952,1033-1344 <lig1></lig1>
A; Cr	oss-references: GB:M32738; GB:U05288; NID:9182121; PIDN:AAA35804.1; PID:9182124; perimental source: solice form a
A; NO	te: GenBank entries HUMEPISIA1 and HUMEPISIA2 present only the amino-and carboxy
A; Mo.	
A;Re: A;Crc	182126 - GIONARAR SEGOR
A; Exi A; Not	or a contract of the contract
R;Gei J. Bj	Gendler, S.J.; Lancaster, C.A.; Taylor-Papadimitriou, J.; Duhig, T.; Peat, N.; Bur.
A; Tit	nan tumor-associated polymorp
A; Acc	
A; Sta	itus: not compared with conceptual translation
A; Res	idues: 1-19, 29-992, 1033-1344 <gen></gen>
A;Crc A;Not	N: AAA59876.1; PID: 9188870
R; Lan J. Bi	3.; Hollingsworth
A, Tit	le: Cloning and sequencing of a human pancreatic tumor mucin cDNA.
A; Acc	
A;Sta A;Mol	tus: not compared with conceptual translation ecule type; mRNA
A;Res A;Cro	1344 <lan></lan>
A; Not	:AAAbUUly.1; PID:g189599 er copies of the tandemly
Eur.	Smorodinsky, N.; Horev, J.; Zaretsky
A;Tit A;Ref	le: Human epithelial tumor antigen cDNA sequences. Differential splicing may gen erence number: s1657: Math. 00356413
A; Acc	ession: S10572
A; Res	ectie type: mkna idues: 1-19,29-155,'P',157-175,'P',177-182,'A',184-212,1033-1344 <wrps< td=""></wrps<>
A;CTO R;Wre.	
submi A;Ref	tted to the EMBL Data Library, March 1990 erence number: \$40293
A; Acc. A; Mole	ession: S40293 ecule type: mRNA
A; Resid	Residues: 1-19,29-155,'P',157-175,'P',177-182,'A',184-212,1033-1037,'A',1039-1344 < Cross-references: EMBL:X52229, NID.,27053, DIDAL,738,5476,1
R; Abe	M.; Siddigui, J.; Kufe, D.

Db 522 PPTPQSVPSAVTEPKPQPRAAPPPSNEATPAVPSPSPNLKSPLPTTPRVPLMALTPQPV 581 Qy 574 TPKETAPTTPEKLAPTTPEKPAPTTPEELAPTTP	Db 582 TAQMYTQLAATKPSPIVPRASPK-ALMTPPPPPGLPRALAAAKLLGLPSSPVASAMHAK 640 Qy 616EEPTPTTPEEPAPTTPRAAAPNTPKEPAPTTPKEP 650 1 1 1 1 1 1 1 1 1 1	651	QY 678 TLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEP 727	728	Qy 781 Db 873	933		KAPKKPTSTKKPKTMP			Qy 1064 GE 1065 Db 1118 RE 1119		C.)Date: 15-Oct.1999 #sequence_revision 15-Oct.1999 #text_change 18-Feb-2000 C;Date: 15-Oct.1999 #sequence_revision 15-Oct.1999 #sequence_revision 15-Oct.1999 #text_change 18-Feb-2000 C;Accession: 725697 R;Fulton, B.	submitted to the Embi Data initial, august 1908. A; Description: The sequence of C. elegans cosmid F16F9. A; Reference number: 220071 A; Accession: T25697	rransiateu >> 3L:U67956; E strain Bri	C;Genetics: A;Gene: CESP:F16F9.2 A;Gene: CESP:F16F9.2 A;Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3 A:Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3	9.3% llarity 26.6% Conservative	
QY 725 -KEPAPTTPKEPAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPT 783	Db 1001 MSSPPPPEWKSPPPPAPVSSPPPPAPVSSPPPPVKSPPPPAPVSSPPPV 1057 ` Qy 784 -TPKETAPTTPKEPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPEL 842 Cy 184 -TPKETAPTTPKEPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPEL 842 Cy 184 -TPKETAPTPKEPAPTSSPPPVKSPPPAPVSSPPPPVKSPPPPAPVSSPPPPAPVS 1115	843	RESULT 8 T18535	high molecular mass nuclear antigen - Chicken (Iragment) C;Species: Gallus gallus (chicken) C;Date: 1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999 C.accession: T18535	R.Shimada, K.; Harata, M.; Mizuno, S. J. Cell Sci. 110, 3031-3041, 1997 A.Title: A nuclear matrix-associated high molecular mass nuclear antigen, HMNA, of chick A.Reference number: Z18955; MUID:9803440	A;Accession: T18335 A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: mRNA A;Residues: 1-1151 <shi></shi>	1151;	324; Ga	OY 67 AEVHNPTSPPSSKKAPPPSGASGTIKSTTKRSPKPPNKKTKKVIESEEITEEH 120	121 SVSENQESSSSSSSSSSSTIWKIKSSKNSAANRE	بە تى	239	297	233 VICHERARE RAITHEANNING VICTOR SAND SAND SAND SAND SAND SAND SAND SAND	DD 291 AINVAREARETSESTERNING TO THE TOTAL OF THE TOTAL OF THE TOTAL OF THE TREPART TOTAL OF THE		478	-

90;

QHYMECCPDFKRVCTAAEVHNPTSP	TKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSS 135	25		PPHHD 29	:	TTKGPALTTPKEPTPT 31	37	TPKEPASTTPKEPTPTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTKEP 366 	APTHPKEDADmmmrgn.commm.commmrs.commrs.commmrs.commmrs.commmrs.commmrs.commmrs.commmrs.commmrs.commrs.commmrs.commmrs.commmrs.commmrs.commmrs.commmrs.commmrs.commrs.commmr		1	TESSSAPVPIPSSSTIESSSAPAPIPSST 534					PEKPAPITPEELADI-TPEEPTPITPEEPAPTIPKAAAPNIPKEPAPITPKEPAP 652	PTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAP 694	TTPKEPAPTTPKETAPT-TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKP 711	VTSSTTESSSAPVPTPSSSTTESSSSAPVPTPSSSTTESSSAPVPTPSSSTTESSS 749	APTTPKGTAPTTPKEPAPTTPKEPAPT-TPKGTAPTTLKEPAPTTPKKPAPKELAPT 767		TTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTP 822	PIPOSSSNITSSAPSSTPFSSSTESSSVPVPTPSSSTTESSSAPVSSSTTESSVAPVPTP 869	TTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKIPAATKPEMTTA 877	TTESFSTGT-TVTPSSSKYPGSQTETSVSSTTETTIVPT 925	KITISVITESVITESVITESVITESVITESVITESVITESV	THIKTHOI ABYUM	TVCSTGINSAGETISGCSPKILITIVPCSTSPEIR. TVCSTGINSAGETISGCSPKILITIVPCSTSPEIR.	TPRKMTSTMPELNPTSPTAFAM	STIVVITEYSTEKBGGEITTTEVTKNIPTTYLTTIAPTPSVITVINFTPIT-ITTTV 1094
23 SSQDLSS-CAGRCGEGYSRDATCNCDYNC 	76 PSSKKAPPPSGASQTIKSTTKRSPKI		136 SSSSTIWKIKSSKNSAANRELQKKLK : : : : 253 SSSTTAPATPTTSCTKFKE		292 TIPCTKKKITISKICTKKTIT	256 KQTSTDGKEKTTSAKETQSIEKTSAK			АРТТРКЕВАВ															FIFSS	SSSSN					TKAPKKPTSTKKPKTM	: : STIVVITEYSTSTKPGGEITTTFVTKNI
Oy Db 1	λo	_	94 Bp 11	0y 15	Dp 25	0y 25	338	Oy 316 Db 380	Oy 367	Db 438	y 423	b 488	y 480	535	539				•		712		768		870	878	926	936	986	987	1038
	•	_	<u> </u>	J		0	Ω	0 0	Ó	Ω	ò	g	ò	g	ò	පු	ð á	ž (õ	g	ò	<u>a</u>	òá	3 8	8 8	ò	o G	ò	: 점	ð	g

	1035 LQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNOG 10 1095 CSTGTNSAGETTSGCSPKYTTTVPCSTGTGEVTTPATTI UTTAVUTTT 1
Qy	1095 IIINPMLSDETNICNGK
qa 	1142 TVVTTESSTGTNSA-GKTTTGYTTKSVPTTYVTTLAPSAPVTPATNAVP 118
Qy	1170
qa	1190 TIITTECSAATNAAGET 1207
RES	RESULT 7
ext	osson - like protein - maize
150	
C; 2	er "cext_cna
sur	nmitted to the EMBL Data Library, June 1994
1 K	Jestiphion: Pex genes: pollen-specific genes with extensin-like domains, efference number: \$49915
A A S	iccession: S49915 Figure: prediminary Oldrula turns: Nun
A;R	A;Residues: 1-1188 <rub> A;Cross-references: EMBL:234465; NID:9600117; PIDN:CAA84230.1; PID:9600118</rub>
0	Score 784: DR 2: Tongth 1100
ω <u>Σ</u>	
٥y	Capa (22) Capacita Arabata Arabata (22)
qa	
QY	28
QΩ	1 1
οy	
qu	
οy	
qq	
Qy	
qa	
ογ	rPK
QQ	
οy	
qa	
Οy	
qa	
δy	67
qq	
δλ	679 LKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTP 724
q	

```
Query Match
Best Local Similarity 27.2%; Pred. No. 2.5e-27;
Matches 332; Conservative 114; Mismatches 519;
                                                                                                                                                                                                       g
                                                                                  qq
                                                                                                                                           g
                                                                                                                                                                            QY
                                                       QΥ
                                                                                                                   δŽ
                                            A;Map position: 10
A;Introns: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3
A;Introns: 24/1: 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3
A;Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding activ
C;Keywords: alternative splicing; DNA binding; transcription factor
                                                                                                                                                                  51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1294 APEETSTTP-----SPOKIPKVAGPKEASATPPSKKTPKTAVFKETSAPSEGVTAVPLE 1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: EMBL:U48363; NID:g1666688; PID:g1666689; PIDN:AAB18732.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             443
                                                                                                                                                                                                                                                                                                                                                                                                                                  ETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKE-PAPTTTKSAPTTPK 356
                                                                                                                                                                                                                                                      SSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTFKPPVVDEAGSGLDN 190
                                                                                                                                                                                             71 NPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSS 130
                                                                                                                                                                                                                          -----PPSPKGAPIVPTESSISSKQVPAEILPSPQ 792
                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKKLTPTTPEKLAPTTPEK-PAPTTPEELAPTTPEEPTPTT-----PEEPAPTTPKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----KKPAPKELAPTTKEPTS-----TTSDKPAPTTPKGTAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAPTTPKE-PAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KETAPTTPKEPAPTTPKKPAPTTPETP ----PPTTSEVSTPTTK -----EPTTIHKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                        998 PEGVTAVPLEIPPCSKKAPKTAAPKESSATSSSKRAPKTAVSKEIPSKGVTAVPLEISLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1058 LKETSKSATPGEKSASSPKRSPKTAGPKE---TPPGGYTAVPPEISLPPKETPQNATPNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SLAASSQKRSPKTSVPKETPPGGVTAMPLEIPSAPQKAPKTAVPKQIPTPEDAVTILAGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       634 AAPNIPKEPAPITPKE-PAPITPKEPAPITPKETAPTIPKGTAPTILKEPAPITD----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----TPKEPAPTTPKK-PAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAPTAPKKPAPTT -- PKEPAP -- -- TTPKEPAPTTTKEPSPTTPK -- -- - EPAPTTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTTTKEPAPTTT-KSAPTTPK-----EPSPTTTKEPAPTTPKEPAPTTPKKPAPTT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  543 -PKE------PAPTTPKEPAP-----TTTKKPAPTAPKEPAPTTPKETAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----EPAPITIKEPAPITPKEPAPITIKEPAPIT--IKSAPI
                                                                                                                                                                                                                                                                                                                191 GDFKVTTPDTSTTQHNK----VSTSPKITTAKPINPRPSLPPNSDTSKETSL--TVNKE
                                                                                                                                                                                                                                                                                                                                           -----DISATLSLKSVPAVTSLSPPKAPVAPSNEATIVPTEIPTSLKNALAAATPKE
                                                                                                                                                                                                                                                                                                                                                                                                    886 TLATSIPKVTSPSPQKTPKSVSLKGAPAMTSKKAT---EIAASKDVSPSQ--FPKEVPLL
                                                                                                                                                                                                                                                                                                                                                                         T-----TVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKA
                                                                                                                                                                  Indels 246;
                                                                                                                                    Length 2187;
                                                                                                                                                                                                                                                                               Best Local Similarity 27.4%; Pred. No. 3.4e-27;
Matches 311; Conservative 128; Mismatches 449;
                                                                                                                                       Score 812; DB 2;
pred. No. 3.4e-27;
                                                                                                                                           11.18; 27.48;
                                                                                                                                                                                                                                 746 SPT-PPSSKGAPVPSTGA--
                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             444
                                                                                                                                                                                                                                                                   131
                                                                                                                                                                                                                                                                                                                                                                                      244
                                                                                                                                                                                                                                                                                                                                                                                                                                               298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        357
                                                                                                                                                                                                                                                                                                                                                        834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΟŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Óγ
                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                    òγ
                                                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                             ò
```

```
RESULT 6
548418
548418
548418
548418
548418
548418
548418
548418
548418
74-alpha-glucosidase (EC 3.2.1.3) - yeast (Saccharomyces cerevisiae)

Nilternate names: extracellular glucoamylase; mucin-like protein MUC1; protein ylecters accharomyces cerevisiae
C.Species: Saccharomyces cerevision 10-Sep-1999 #text_change 12-Nov-1999
C.Accession: 548478; A26877; B26877; S27281; J66123
R; Rowley, K.
submitted to the EMBL Data Library, October 1994
A; Recession: 548478
A; Accession: 548478
A; Accession: 548478
A; Accession: 548478
A; Residues: 1-1367 < ROW
A; Residues: 1-242 < RAM
A; Residues: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..,..r.r.r.c.i.v.r.y.c.superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosida C;Superfamily: yeast glucan 1,4-alpha-glucosida C;Keywords: glycosidase; hydrolase; polysaccharide degradation; transmembrane protein F;5-21,/Domain: transmembrane #status predicted <TM1>
F;5-21,/Domain: transmembrane #status predicted <TM2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         psend
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         is critical for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Cross_references: GB:U30626; NID:g1304386; PIDN:AAC49609.1; PID:g1304387
                                            A)ross-references: EMBL:X13857; NID:94551; PIDN:CAA32069.1; PID:94552 A)rCross-references: EMBL:X13857; NID:94551; PIDN:CAA32069.1; PID:94552 R:Lambrechts, M.G.; Bauer, F.F.; Marmur, J.; Pretorius, I.S. Proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, 1996 Affittle: Mucl., a mucin-like protein that is regulated by Mss10, is crit A; Reference number: JG6123; MuID:96323237
                                                                                                                                                                                                                                                                                                                                     ITTTEIMNKPEETAKPKDRATNSKATTPKPOKPTKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                                                                                                                                                                                                                                                                                                                      PDESTPELSAEPTPKALENSPKE-PGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETT
                                                                                                                                                                  894 TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKT
                                                                                                                                                                                                                       1014 PRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETP 1067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Gene: SGD:MUC1; STA2; MAL5; DEX2; SGD:S0001458
A;Cross-references: MIPS:YIR019c; SGD:S0001458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Residues: 1-1367 < LAM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-31 <PAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: S27281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1691
                                                                                                                  1580
                                                                                                                                                                                                                                                                                                                                                                        954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1731
                                       835
```

54;

Gaps

255;

Length Indels

```
.996
ing-in of a proline-rich exon converts alphaNAC into a mus
MUID:96312450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ated complex alpha chain, muscle splice form gp220 - mouse AC protein
buse mouse)
nce_revision 22-Oct-1999 #text_change 24-Nov-1999
                TTKEPAPT-----TTKSAPTTPKEP----APTTPKKPAPTTP 410
PITPKEPAPITIKSAPITPKEPAPITIKEPA------367
                                                                 TPTTPKEPAPTTKEPAPTTPK-EPAPTAPKKPAP-----TTPKE 459
                                                                                                                     EP----SPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTP-KEP 514
                                                                                                                                                                          ВРАРТТРККРАР - - ТТРКЕРАРТТРКЕРАРТТККРАРТА-РКЕ 569
                                                                                                                                                                                                                                                                            KDOKLKPSPAKKPEKAPEPAAPKKWKPVWDDDPDEPEADFTVPA 1170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 619
                                                                                                                                                                                                                                                                                                                               PNTPKEPAPTTPKEPAPTT-PKEPAPTTPKETAPTTPKGTAPTT 678
                                                                                                                                                                                                                                                                                                                                                                                                                                       LAPT-TTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAP 737
                                                                                                                                                                                                                                                                                                                                                                                   T-----TPKKPAP-----KELAPTTTKGPTS-----TTSDK 779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----APTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTKEP 828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ELSA-EP---TPKALENSPKEPGV-------PTTKTPA 867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KDKTTERDLRTTPETT---TAAPKMTKETATTTEKTTESKITATT 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKAT 979
                                                                                                                                                                                                                                                           KLTPTTP----EKLAPTTPEKPAPTTPEELAPTTPEEP----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KKPKTMPRVRKPK------TTPTPRKMTSTMP 1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     slated from GB/EMBL/DDBJ
```

m

RESULT T31108

cyst germination specific acidic repeat protein precursor - Phytophthora infestans (; page 22-oct-1999 #sequence_revision 22-oct-1999 #text_change 22-oct-1999 #sequence_revision 22-oct-1999 #text_change 22-oct-1999 (; pate: 22-oct-1999 #sequence_revision 22-oct-1999 #text_change 22-oct-1999 (; Accession: T31108 A; Reference number: 220986 A; Reference number: 220986 A; Reference number: 220986 A; Reference number: 220986 A; Residues: preliminary; translated from GB/EMBL/DDBJ A; Atatus: preliminary; translated from GB/EMBL/DDBJ A; Cross-references: EMBL:AF061185; NID:93851513; PID:93851514; PIDN:AAC72308.1 A; Cross-references: EMBL:AF061185; NID:93851513; PID:93851514; PIDN:AAC72308.1 A; Genetics: 51; 1047 548 809 909 PEELAPT--TPEEPTPTTPEEP-----APT-----TPKAAAPNTPKEPAPTTPK 648 927 499 749 462 689 425 629 258 454 318 511 375 570 394 TKKVIESEEITEEHSVSENQESSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNK 166 334 -TIKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEP-----APTTPKGTAPT-------TLKEPAPTTPKKP-----APKE---LAPT--TTKGPTSTTSDKPAP EP----APTTPKEPAPTTPKETAPTTPKGTAPT--TLKEPAPTTPKKPAPK-ELAPT-TTPKEPAPT--TTKEPSPTTPKEP-----APT-TTKSAPT-----TTKEP---T--TPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTT 630 TYAPTEETTYASTEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAP PKEP-----APTAPKKPAPTTPKEPAPTTPKEP-----APTAPKKPAPTTPKEPAP KNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTT-----QHNKVSTSPKITTAKPI STDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTPK 455 YAPTEETTYAPTEETTYAPTKETTYAPT---EETTYASTEETTYAPTEETTY EPASTIPKEPIPITIKSAPITPKEPAPI-TIKSAPITPKEPAPI--TIKEPAPITPKEPA 376 PITIKEPAPITIKSAPI------TPKEPAPITPKKPAPITPKEPAPITPKEPIPIT -----PNSDT----SKETSLTVNKETTVETKETTT---TNKQT ::| | ::| |:::| |:::| |:::| |:::| |:::| |::::| |::::| |::::| |:::| |::::| |::::| |::::| |::::| |::::| |::::| | 229; Length 1489 Indels 11.8%; Score 863.5; DB 2; 31.7%; Pred. No. 1.5e-29; iive 55; Mismatches 509; Conservative Local Similarity es 368; Conserv NPRPSLP-Query Match Best Local S Matches 368, 868 649 549 209 928 463 319 107 167 221 395 259 qq g οy qq QΥ qq Ωÿ qq δ ò g ДQ δ g QΥ δy ò g g οy q ΩÝ Dp ò δy

```
Slayer protein - Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Date: 15-0ct-1999 #sequence_revision 15-oct-1999 #text_change 15-oct-1999
C;Accession: T18262
R;Fujino, T; Beguin, P; Aubert, J.P.
J. Bacteriol: 175, 1891-1899, 1993
A;Title: Organization of a Clostridium thermocellum gene cluster encoding the
                                                       1683 SPITTTP--SPPTT----TMTTPS--PTTTPSSPITTTTPSSTTTTP-SPPTTMTTPSP 1733
                                                                                                                                                                                                 1794 SÖKPNFHKPGGDTELIGDVCGPGWAANISCRATMYPDVPIGQLGQTVVCDVSVGLICKNE 1853
                                                                                                                                                                                                                                                                                                                                                                                                            TTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTP----KETAPTTP----- 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1970 TVLITTTTMIPIPISTKSTTVIPITTTTTTVTATPTPTGTOTPTMIPISTTTTVTPTP 2029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------TAPTTLKEPAPTTPKKPAPKELAPTT---TKEPTS 705
                                                                                                                                                                                                                                                                                                                                                                                      TISDKP----APT-TPKGTAPTIPKEPAPTTPKEPAPTIPKGTAPTTLKEPAPTTPKKPA 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----TAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPR 1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2261 -----QENETWWLCDCFM-ATCKYNNTVELVKVECEPPPMPTCSNGLQPVRVEDPDGC 2312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1062 ------AEGETPHMLL------RPHVFMPEVTPDMDYLPRVPNQGIIIN 1098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2313 CWHWECDCYCTGWGD-PHYVTFDGLYYSYQGNCTYVLVEEISPSVD-----NFGVYID 2364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: EMBL:X67506; NID:9296879; PID:9296881; PIDN:CAA47841.1
                              509 TTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPK
                                                                                                    EPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAP
                                                                                                                                                                                                                                                                                                                                                PKELAPTTTKGPTST-TSDKPAPTTPKETAPTTPKEPAPTTPKKP-----APTTPETP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                813 PPTTSEVSTPTTTK----EPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATKPEMITTAKDKTTERDLRITPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2148 WIPTPIPISIPSIIRITGLRPYPSSVLICCVLNDIYYAPGEEVYNGTYGDICYFVNCSLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2208 CTLEFYNWSCPSTPSPTPTPSK-STPTPSKPSSTPSKPTPGTKPPECPDFDPPR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1016 KMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNP------KSEDAGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       928 QDTTPFKITT---LKTTTLAPKVTTTKKTITTTEIMNKPEE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 950; DB 2;
Pred. No. 3.8e-33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Accession: T18262
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-1664 <PUJ>
                                                                                                                                                                                                                                               KG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: 218847; MUID:93209931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13.0%;
31.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                        904
                                                                                                                                                                          629
                                                                                                                                                                                                                                               672
                                                                                                                                                                                                                                                                                                                 674
                                                                                                                                                                                                                                                                                                                                                                                                                                                            761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             996
                                                          g
                                                                                                                                 qq
                                ò
                                                                                                  Óγ
                                                                                                                                                                     ò
                                                                                                                                                                                                     g
                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                           Ωp
                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
J. Clin. Invest. 87, 77-82, 1991

A; Title: Human bronchus and intestine express the same mucin gene.
A; Reference number: A61257; MUD:91086481
A; Accession: A61257; MUD:91086481
A; Accession: A61257; MUD:91086481
A; Status: not compared with conceptual translation
A; Molecule type: mRnA
A; Residues: 'T',1925-1948, 'TTS',1952-1954 <7AN>
A; Residues: Biophys. Res. Commun. 138, 821-828, 1992
A; Title: Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-t A; Molecule type: mRNA
A; Residues: 2328-2468 <XUG>
A; Molecule type: mRNA
A; Residues: 2328-2468 <XUG>
A; Cross-references: GB:M86523
A; Residues: 2328-2342, 'K', 2344-2354 <XUGI>
A; Residues: 2328-3342, 'K', 2344-2354 <XUGI>
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics: GB:MUC2
A; Mad Dossition: 11015, 5-11015, 5
                                                                                                                                                             D.; Wang, D.; Jones, C.; Forstner
                                                                                                                                                                                                       Ü
                                                                                                                                                                                                                                                                                                                                                                                                                                                    type A repeat homology; von
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1079 SCS--CDIG--GDCECFCSAVASYAQECTKEGACVFWRTPDLCPIF-----CDYNNPH 1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1129 ECEWHYEPCGURSFETCRTINGIHSNISVSYLEGCYPRCPKDRPIYEEDLKKCVTADKCG 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              75 PPSSKKAPPPSGASQIIKSTT------115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1326 FDGVCGAPEDIECRSVKDPHLSLEQHGQKVQCDVSVGFICKNEDQFGNGPFGLCYDYKIR 1385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1386 VNCCWPMDKCITTPSPPTTTTTTTTTTPSPPTTTTTTTTTTTTTPPPTTTTPSPPTTTT 1445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --ITEEH----SVSENQESSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNR 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29 SCAGRCGEGYSRDATCNCDYNCQHYMEC------CPDFKRVCTAAEVHNPTS 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 333 IKSAP-TTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAP 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -TTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTTPKEPAPTTKEPAPTTPKEPAPTAPK 450
                                                                                                                                                                                                                                                                                                                                                                                                                  Map position: 11p15.5-11p15.5
Superfamily: von Willebrand factor; von Willebrand factor type A representation of Marchelin; intestine; tandem repeat
2766-2834/Domain: von Willebrand factor type C repeat homology <VWC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            170 TKKKPTPKPPVVDEAGSGLDNGDF···KVTTPDTSTTQH···NKVSTSPKITTAKPINPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PSLPPNSDTSKETSLTVNKETTVETKET--TTTNK-------QTSTDGKEKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----LAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAP-TTTKEPAPTTTKSAP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
14.7%; Score 1073; DB 2; Length 3020;
Best Local Similarity 27.1%; Pred. No. 4.3e.38;
Matches 374; Conservative 101; Mismatches 502; Indels 402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     267 TS--AKETQSIEKTSAKD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
```

2207

867

cellulo

```
Compugen Ltd
GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
```

- protein search, using sw model OM protein

(without alignments)
907.900 Million cell updates/sec April 26, 2002, 16:31:52 ; Search time 114.61 Seconds Run on:

1 MAWKTLPIYLLLLSVFVIQ.......ARAITTRSGQTLSKVWYNCP 1366 **BLOSUM62** Title: Perfect score: Sequence: Scoring table:

219241 seqs, 76174552 residues Searched:

Gapop 10.0 , Gapext 0.5

219241 hits satisfying chosen parameters: Total number of

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

pirl:* pir2:* pir3:* PIR_68:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description		S-layer protein -	u			glucan I,4-alpha-g		ro .	hypothetical profe	precurso	i.		v	nucleolar phosphop	extensin class 1 p	cal prot	α	process one of			hypothetical proce		larval glue protei	hypothetical proce	ascites slalogifico			hypothetical prote	_	njurofilament trin	1
	ID	A43932	T182					S49	T18			138												Ų1	T34	Ø	S22	U	3 6	- '	2067	OFMSH
	Length DB	3020 2		_	1274 2																			379 2								1087 1
dР	Query Match Le	1.	* C		11.6		-	-	6	σ	σ	. 0	ά				ò	ю ю	œ		8	ω	α	œ	00	7		. 1	. 1	7	7.6	7.6
	Score	1073	010	0000	846	212	21018	784	100	0 8 9	629	1 1 2 2				643.0	633	632	629	627.5	622.5	607.5	700	60.00	592.5	α . α . α . α . α . α . α . α . α . α .	0 0 0 0	, ,	600	556.5	551	
	Result No.		٦ (7 0	n «	* 4	י ני	0 1	- α	0	,	9 -	1.	77	13	T 4	12	16	17	18	6	6	3 5	22	3 6	3 6	* u	7 (26	27	28	29

A;Cross-references: GB:M22405; NID:g188873; PIDN:AAA36334.1; PID:g188874 A;Experimental source: intestine R;Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.; Basbaum, C.B.

A, Molecule type: mRNA A, Residues: 1916-2193 <GU4>

A; Accession: B33532

calphotin - fruit hydroxyproline-ric extensin precursor hypothetical prote neurofilament trip hypothetical prote vsaA protein precu exo-alpha-sialidas calcium-binding pr tegument protein 2 hydroxyproline-ric neurofilament trip membrane glycoprot hypothetical prote murin-like glycopr	membrane glycoprot
96233572752111503	ოლ
A47283 S20500 JU00465 T32271 A37221 T27642 S70795 A47282 T42567 T42567 T42567 T42567 T42567	T3111 T4546
000000000000000000000000000000000000000	2 0
873 873 416 416 1459 1075 813 1162 865 3534 3534 865 866 811	1832
77777777777777777777777777777777777777	7.0
545 542 538.5 537.5 537.5 530.5 530.5 520.5 521.5 521.5	513 512.5
0 1 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2	44

ALIGNMENTS

```
A;Status: not compared with conceptual translation
A;Status: not compared with conceptual translation
A;Rolecule type: mRNA
A;Rolecule type: mRNA
A;Roresidues: 2037-3020 <GU3>
A;Cross references: GB:M94132; NID:g186397; PIDN:AAA59164.1; PID:g186398
A;Cross references: GB:M94132; NID:g186397; PIDN:AAA59164.1; PID:g186398
A;Cross reference color
A;Rorest source color
B;Note: Sequence extracted from NCBI backbone (NCBIP:116698)
A;Note: Sequence extracted from NCBI backbone (NCBIP:116698)
A;Note: Sequence extracted from NCBI backbone (NCBIP:116698)
A;Title: MUC-2 human snall intestinal mucin gene structure. Repeated arrays and polym A;Reference number: A43932; MUID:91358717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: DNA A; Residues: 1343-1350, L',1352-1411,'S',1413-1448,'P',1450-1503,'T',1505-1915 <TOR>
A; Residues: 1343-1350, L',1352-1411,'S',1413-1448,'P',1450-1503,'T',1505-1915 <TOR>
A; Cross-references: GB:M74027; NID:9188863; PIDN:AAA59875.1; PID:9188864
A; Cross-references: GB:M74027; NID:918864
A; Note: sequence inconsistent with the nucleotide translation
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence inconsistent with the nucleotide translation of human intestinal mucin cDNAs. Sequence analysis and eving the number: A33532; MUID:89197956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: mRNA
A; Molecule type: mRNA
A; Residues: 1-639 (GU1>
A; Residues: 1-639 (GU1>
A; Cross-references: GB:L21998
A; Cross-references: GB:L315-23183, 1992
J. Biol. Chem. 267, 2135-23183, 1992
A; Title: The human MUC2 intestinal mucin has cysteine-rich subdomains located both up A; Reference number: A45106; MUID:93016075
A; Accession: A45106
mucin 2 precursor, intestinal - human (fragments)

N.Alternate names: mucin SNUC-41

C.Species: Homo sapiens (man)

C.Date: 10-Mar-1993 #sequence_revision 12-Apr-1996 #text_change 05-Nov-1999

C.Date: 10-Mar-1993 #sequence_revision 12-Apr-1996 #text_change 05-Nov-1999

C.Accession: A49963; A45106; B45106; A43332; B33532; A61257; PQ0328; PQ0329

R.Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim, Y.S.

J. Biol. Chem. 269, 2440-2446, 1994

A;Title: Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the Accession and the A49663; MUID:94132002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: mRNA
A; Residues: 626-1895 <GU2>
A; Residues: 626-1895 <GU2>
A; Cross-references: GB:M94131; NID:9186395; PIDN:AAA59163.1; PID:9186396
A; Cross-references: GB:M94131; NID:9186395; PIDN:AAA59163.1; PID:9186396
A; Accession: B45106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A, Status: not compared with conceptual translation
```

882	N a	. 0	œ		1058	2 _					
837 TTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTKTPAA 882 1 :	883 -TKPEMITTAKDKTTERDLRITPETITAAPKWIKETATTEKTIESKITAATHUNUTC 028	1	939 TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPO 998	821 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXP- 868		869 -GSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSA-PDXRPXPGST 911	-PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNOGTITNBMLSD	912 APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHG 957	ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPSSPSPARRITERWACTEREDIT	1	
9.5	88	16	93	82	66	98	1059	91	1118	95	
g g	ò	qq	ò	g	δy	QQ	ò	qq	δ	qq	

Search completed: April 26, 2002, 16:16:52 Job time: 203 sec ACC,

```
= CCT, CCC.
                                                                /note= "Amino acid 144 is Y = Xaa
which is the codon for Thr or Asn wherein Thr = ACT,
or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                               53;
909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TTP------KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP 746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    123 -- DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            250 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 308
                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                       Pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                586 PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KPAP--TTPK----ETAPTTPKEPAPTTP------KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --KKPAPKELAPTTTKGPTSTTSD
 or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       536 PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     588 TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      420 SAPDXRPXPGSTAPXAHGVTS----PDDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393 TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                      16 TVLTV------VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    309 P-KAETTTKGP--ALTTPKEP-----TPTTPKEPASTTPKEPTTIKSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --APTTTKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  440 PKE--PAPTTKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTT
                                                                                                                                                                   /note= "Amino acid 147 is X2 = Xaa
which is the codon for Pro or Ala wherein
or CCG; and Ala = GCT, GCC, GCA, or GCG."
 GCA,
                                                                                                                                                                                                                                                                                                                                                                                                 241;
                                                                                                                                                                                                                                                                                                                                                            Length 2035;
 = GCT, GCC,
                                                                                                                                                                                                                                                                                                                                                              5.7%; Score 424.5; DB 4; Length 2 ilarity 23.7%; Pred. No. 2.4e-18; Conservative 69; Mismatches 511; Indels
                                                                                                                                                                                                                                                                            /note= "Amino acids 1 to 21 are a 21 amino acid precursor sequence.
   CCG; and Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        747 KEPAPTTPKG----TAPTTLKEPAPTTP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTPKEPAPTTT-----KSAPTTPKEP--
         or
         CCA,
         ccc,
                                                                                                                                                                                                                                                  NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION:
OTHER INFORMATION:
9-083-116-2
                                                                                                                                               NAME/KEY: Peptide
LOCATION: 147
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                          LOCATION: 144
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
       OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                             NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                            255;
                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                    FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         644
                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                           -60-SN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              \Omega Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               repeat .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       = CCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Amino acid 134 is X1 = Xaa
Xaa Xaa which is the codon for Pro or Ala wherein Pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "The amino acids spanning
128 to 1899 constitute a repeated region wherein the
20 amino acids, 17 of which are fixed. The number of
repeats varies from 1 to 40."
                  ---GVTSAPDXRPXPGSTA----PXAHG 957
                                                   1118 ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPID 1172
                                                                           GENERAL INFORMATION:
APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: LATHE, R. Mara
APPLICANT: HARBUYENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
                                                                                                                                                                                                                                                                                                                                                                    F.O. BOX 1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO PCT/FR91/00835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PFILING APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00
FILING DATE: 23-0CT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US/09/083,116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08/479,537
                                                                                                                                                                              Sequence 2, Application US/09083116
Patent No. 6203795
                         912 APXAHGVTSAPDXRPXPGSTAPXAH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2035 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: Virginia COUNTRY: United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Peptide
LOCATION: 128..1899
OTHER INFORMATION: /r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY: Peptide LOCATION: 134 OTHER INFORMATION: OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                          CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                              22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE
                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                     US-09-083-116-2
                                                                                                                                                                                                                                                                                                                                                                                              STREET:
                                 g
                                                                   δý
                                                                                                   g
```

53;

```
250 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 308
                                                                                                                                                                                                                                      PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP 122
                                                                                                                                                                                                                                                                                352 TIPKEPAPTIT-----KSAPTIPKEP------APTITKEPAPTIPKEPAPT 392
                                                                                                                                                                                                                                                                                                                                                                                  351
                                                                                                                                                                                                                                                                                                                                                                                                                                                486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          487 TKEPSPTTFKEPAPTT--TKSAPTTTKEPAPTT-----TKSAPTTPKEPSPTTTK----E 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    644 TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTLKEPAP 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              420 SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD 1117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | : : : | | : : : | | DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               821 APDXRPXPG----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP- 868
                                                                                                                                                        309 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTTIKSAP
                                                                                                                                                                                                                                                                                                          393 TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                  PKE--PAPTIKEP-----APTIPKEPAPTAPK----KPAPTIPKEPAPTIPKEPAPTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301 APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  536 PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       747 KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       472 APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTP-----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         645 FPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   883 -TKPEMTTTAKDK---TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TPKETAP ----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               793 KPAP--TTPK----ETAPTTPKEPAPTTP------KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       837 TTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               763 DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIABAMLQTTTRPNQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241;
                                                                  Length 2035;
                                                                                                    Indels
                                                            Score 424.5; DB 2;
Pred. No. 2.4e-18;
); Mismatches 511;
                                          5.7%; Scc..
23.7%; Pred
                                                                        Best Local Similarity 23.77
Matches 255; Conservative
US-08-479-537A-2
                                                          Query Match
                                                                                                                                                                                                                                            99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         666
                                                                                                                                ð
                                                                                                                                                                  q
                                                                                                                                                                                                        à
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "Amino acid 134 is X1 = Xaa
Xaa Xaa which is the codon for Pro or Ala wherein Pro = CCT,
CCC, CCA, or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note- "The amino acids spanning 128 to 1899 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "Amino acid 144 is Y = Xaa
which is the codon for Thr or Asn wherein Thr = ACT, ACC,
or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ccc,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note- "Amino acid 147 is X2 - Xaa which is the codon for Pro or Ala wherein Pro - CCT, or CCG; and Ala - GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "Amino acids 1 to 21 are a
21 amino acid precursor sequence."
CORRESPONDENCE ADDRESS:
ADDRESSREE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. BOX 1404
CITY: Alexandria
                                                                                                                                                                                                  SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILLIG DATE: 07-JUN-1995
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
FILING DATE: 23-0CT-1990
PRIOR APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 33-0CT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 14-APR-1993
FILING DATE: 14-APR-1995
ATMOREY/ABENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: TESKIN, ROBIN L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELEPHONICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2035 amino acids
                                                                  STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: Peptide
LOCATION: 128..1899
OTHER INFORMATION: //O
OTHER INFORMATION: 12
OTHER INFORMATION: re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION: 1..21
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LOCATION: 134
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURE:
NAME/KEY: Peptide
LOCATION: 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: amino a STRANDEDNESS;
                                                                                                                                              MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURE:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURE:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH:
```

```
999 KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT 1058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .059 -PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD 1117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     939 TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRAINSKATTPKPQ 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               702
                                                                                                                                                                                                                                                                                                                                        471
                                                                                                                                                                                                                            587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1118 ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSFARRITEVWGIPSPID 1172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HARBUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             912 APXAHGVTSAPDXRPXPGSTAPXAH------GVTSAPDXRPXPGSTA----PXAHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           645 RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DXRPXPGSTAPXAHGVTSAPDXRPXAPGSTAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               883 -TKPEMTITAKDK---TTERDLRITPETITAAPKMIKETAITTEKTIESKITAITTQVIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 763 DXRPXPGSTAPXAHGVTSAPDXRPXPGST -- APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           821 APDXRPXPG----STAPXAHGVT----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  586 PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX
                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----KKPAPKELAPTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  747 KEPAPITPKG----TAPITLKEPAPITP------KKPAPKELAPITKGPISITSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     837 TITKEPIT-----IHKSPDESTPELSAEPIPKALENSPKEPGVPITKIPAA-----
                                                                                                                                                                                                                                                                                                                                      SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                              TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                  472 APXAHGVISAPDXRPXPGSTAPXAHGVISAPDXRP-XPGSTAPXAHGVISAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KPAP--TTPK----ETAPTTPKEPAPTTP------KKPAPTTPETPPPTTSEVSTP
                          187 TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                      TPKETAP----TTPKKLIPTTPEKLAPITPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                          ---APITPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                            241 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                      301 APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                            536 PAPITPKEPAPITPK----KPAPITPKEPAPITPK----EPAPITIKKPAPIAPKEPAPI
                                                                                                                                                                                                                                                              SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     869 -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPX--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08479537A Patent No. 5861381 GENERAL INFORMATION:
                                                                                    PKE--PAPTTKEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
US-08-479-537A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---- TLB 669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       703
                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                  588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΟŊ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΟŸ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ÓΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                Pb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ΩŽ
                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                            셤
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                            δ
                                                g
                                                                                        δ
                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "The amino acids spanning
128 to 1727 constitute a repeated region wherein the repeat
20 amino acids, 17 of which are fixed. The number of such
repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Amino acid 144 is Y = Xaa which is the codon for Thr or Asn wherein Thr = ACT, ACC, or ACG, and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CCI, CCC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Amino acid 134 is X1 = Xaa
which is the codon for Pro or Ala wherein Pro = CCT, CCC,
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----APTTTKEPAPTTPKEPAPTT 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  123 --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                250 ISLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----TPTTPKEPASTTPKEPTPTTIKSAP 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Amino acid 147 is X2 = Xaa
which is the codon for Pro or Ala wherein
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 4; Length 1867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Amino acids 1 to 21 are a
21 amino acid precursor sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5.7%; Score 424.5; DB 4;
ilarity 23.7%; Pred. No. 2.1e-18;
Conservative 69; Mismatches 511;
                                                              WO PCT/FR91/00835
                                                                                                                                                                                                                                      NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION:
TELEPHONE: (703) 836-6620
                                                                                                                                                                      APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                              PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          352 TTPKEPAPTTT-----KSAPTTPKEP-
        08/479,537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  309 P-KAETITKGP--ALTIPKEP--
                                                                                                                                                                                                                                                                                                              TELEPHONE: (703) 836-6620
TELEPAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 amino acids
                                          PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PC
FILING DATE: 23-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                peptide
                                                                                                                                                                                                                                                                                                                                                                                                                              single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME/KEY: Peptide
LOCATION: 128..1727
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: 144
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COCATION: 1..21
CTHER INFORMATION:
CTHER INFORMATION:
US-09-083-116-5
      APPLICATION NUMBER: FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION: 147
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Peptide
                                                                                                                                                                                                                                                                                                                                                                                                              amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
Matches 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99
```

g

Op οy Q

òλ

```
644 TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP 698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         999 KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT 1058
                                                                                                                                                                            747 KEPAPTTPKG----TAPTTLKEPAPTTP------KKPAPKELAPTTTKGPTSTTSD 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1059 •PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD 1117
                                                                                                                                                                                                                                                                            644
                                                                                                                                                                                                                                                                                               793 KPAP--TTPK----ETAPTTPKEPAPTTP-----KKPAPTTPETPPPTTSEVSTP 836
                                                                                                                                                                                                                                                                                                                                645 RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP 702
                                                                                                                                                                                                                                                                                                                                                                                                    DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTQDTTPFKLTTLKTTTLAPKVTTTKKTLTTTEIMNKPEETAKPKDRATNSKATTPKPQ 998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      821 APDXRPXPG-----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP- 868
                                                                                                              472 APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                 699 TTP-----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                                                                    586 PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXRPGSTAP-XAHGVTSAPDX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                             883 -TKPEWTTTAKDK---TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTOVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPID 1172
                                                                                                                                                                                                                                                                                                                                                                     837 TTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPX-------PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            912 APXAHGVISAPDXRPXPGSTAPXAH------GVTSAPDXRPXPGSTA----PXAHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 5, Application US/09083116
Patent No. 6203795
GENERAL INFORMATION
APPLICANT: CHAMBON, Pierre
APPLICANT: LATHE, Richard
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Marie-Paule
APPLICANT: HAREUVENI, Marie-Paule
APPLICANT: HAREUVENI, Marie-Paule
TITLE OF INVENTION: TREATMENT OR PREVENTION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, OBANE, SWECKER & MATHIS, L.L.P.
CTTV. Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER REARABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/NG-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CITY: Alexandra STATE: Virginia COUNTRY: United States 'IP: 22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
US-09-083-116-5
                                          420
                                                                                                                                                                                                                                                                                                                                                                                                            703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1118
                                                                           Qγ
                                                                                                              g
                                                                                                                                               ογ
                                                                                                                                                                                    g
                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                    CCP
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ີວິວ
                                                                                                                                                                                                               region wherein the repear
                                                                                                                                                                                                                                                                                                               /note= "Amino acid 134 is XI = Xaa
which is the codon for Pro or Ala wherein Pro = CCT, CCC,
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "Amino acid 144 is Y = Xaa which is the codon for Thr or Asn wherein Thr = ACT, ACC, or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "Amino acid 147 is X2 = Xaa
which is the codon for Pro or Ala wherein Pro = CCT, CCC,
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   250 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETOSIEKTSAKD-LAPTSKVLAKPT 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           309 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTIKSAP 351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6 PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   352 TTPKEPAPTTT-----KSAPTTPKEP------APTTTKEPAPTTFKEPAPTT 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              123 -- DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393 TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPT 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 440 PKE--PAPTTKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                487 TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTK----E 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301 APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                  /note= "The amino acids spanning
128 to 1727 constitute a repeated regior
20 amino acids, 17 of which are fixed.
repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 2; Length 1867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Amino acids 1 to 21 are a
21 amino acid precursor sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.7%; Score 424.5; DB 2;
23.7%; Pred. No. 2.1e-18;
.1ve 69; Mismatches 511;
   SEQUENCE CHARACTERISTICS:
LENGTH: 1867 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 23.79
Matches 255, Conservative
                                                                         single
                                                                                       TOPOLOGY: linear
MOLECULE TYPE: peptide
                                                                                                                                           NAME/KEY: Peptide
LOCATION: 128..1727
OTHER INFORMATION: /r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCATION: 1.21

COTHER INFORMATION: 7

COTHER INFORMATION: 2:

US-08-479-537A-5
                                                                                                                                                                                                                                                                                                                                                                             NAME.KEY: Peptide
LOCATION: 144
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
FRATURE:
NAME/KEY: Peptide
                                                                                                                                                                                                                     OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                          NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                               LOCATION: 134
OTHER INFORMATION:
                                                  TYPE: amino acid
STRANDEDNESS: sin
                                                                                                                                                                                                                                                                                                                            OTHER INFORMATION:
OTHER INFORMATION:
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
INFORMATION FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                             LOCATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              유
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

ò

```
TELEPHONE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
US-08-479-537A-5
                                                               849
                                                                                                                                                            782
                                                                                             qq
                                                                                                                                                                                           δy
                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                               ŏ
                                                                                                                                                              pp
                                  a
                                                                  ŏλ
                                                                                                                                                                                                                                                                                                           52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     392 SSSSASPFPAS--QALPSYPHSFPPPTSL----SVSNQP----PKYTQPSLPSQAVWSQ 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441 GPPPPPYGRLLANSNAHPGPFPPSTGAOSTAHPPVSTHHHHUQQQQQQQQQQQQQQHHG 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --TPEEPAP------APTTPKAAAPN----TPKEPAPTTPKEP-----APTTPKE 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   281 GGGNLPSAPPPANFPHVTPNLPPPALRPLNNASASP----PGLGAQPLPGHL--PSPYA 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTP--KEPASTTP 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIYSPGSVENDS------DSSSGLSQGPA--RPYHPPPLFPPSPQPPDSTP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 340 KEPTPTIIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                400 --TTKSAPTTPK------EPAPTTPKKPAPTTPKEPAPTTPKEPAPTTPK 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            442 EPAPTIKEPAP-----TIPKEPAPTA--PKKPAPTIPKEPAPTIPKEPAPTITKEPSPTT 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83 ETNAPKKTKTEGELPRPQSPSDLDGENGTSL--NDDGSSDPRDIDQDNRSTS-----P 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222 VSTSPKITTAKPINPRPSLPPNSDTSKETSLTVVKETTVETKETTTTNKQTSTDGKEKTT 281
                                                                                                                                                                                                                                                                                                                                          102 SQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSS 161
                                                                                                                                                                                                                                                                                                                                                                                                       162 KNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP---TTTKEPTSTTSDKPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   501 NSGPPPPGAFPHPLEGGSSHHAHPYAMSPSLGSLRPYPPGPAHLPPPHSQVSYSQAGPNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PPVSSSSNSSSSTSQGSYPCS--HPSPSQGPQGAPYPFPPVPTVTTSSATLSTVIATVAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         495 PKE-----PAPTTTKSAPTTTKEP--APTTTKSAPTTP-KEPSPTTTKEPAPTTPKEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        546 РТТРККРАРТТРКЕРАРТТРКЕ-РАРТТККРАРТАРКЕРАРТТРКЕТАРТТРККLT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 177 RQPEASF------EPHPSVTPTGXHAPMEP-PTSRMFQAP--PGAPPPHPQLXPGGT
                                                                                                                                                                                                                                                                                                                                                                          9 SMSMRSGRKKEAPGPR-----KAEKS
                                                                                                                                                                                                                                                                                                                                                                                                                              303;
                                                                                                                                                                                                                                                                                Length 1185;
                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                    DB 4;
                                                                                                                                                                                                                                                                                Similarity 21.8%; Score 425.5; DB 4; Similarity 21.8%; Pred. No. 1.1e-18; t0; Conservative 119; Mismatches 437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --PTTPE--KLAPTTPEKPAPTTPEELAPTTPEEPTPT--
                                                                          P-LJ 2626
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31.815
REFERENCE/DOCKET NUMBER: 9-LJ
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEPHONE: (619) 535-9001
TELEPHONE: (619) 535-901
TELEPHONE: 1019 535-901
TELEPHONE: 1019 535-901
TELEPHONE: 31.955-901
TELEPHONE: 31.955-901
TELEPHONE: 31.955-901
TELEPHONE: 31.955-901
TELEPHONE: 31.955-901
TENCHMENTION: 31.955-901
TOPOLOGY: 1100-001
TOPOLOGY: 1100-001
                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                          , MOLECULE TYPE:
US-09-041-886-23
                                                                                                                                                                                                                                                                                                                      Matches 240;
                                                                                                                                                                                                                                                                                          Query Match
Best Local 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               729
                                                                                                                                                                                                                                                                                                                                                                                                                                                     51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  óλ
                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
1024 TIPIPRKMISTMPELNPISRIAEAMLQTITRPNQTPNSKLVEVNPKSEDAGGAEGETPHM 1083
                                                                                                                                                                                                                                                                                       964 TKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPK 1023
                                                                                                                                                                                                                                                                                                                                                                                                                       789 TISDKPAPITPKETAPITPKEPAPITPKKPAPITPETPPPTISEVSTPTITKEPTIIHKS 848
                                                                                                                                905 ETTTAAPKMTKETATTTEKT-TESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTT 963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: CHAMBON, Pierre
APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: PREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                           671 -TSPPAGPGTFKPGSPTV--GPGPLPPAGPS-GLPSLPPPPAAPASGPPLS--ATQIKQE
                                                                                              PDE--STPELSAEPTPKALENS--PKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTP
                                                                                                                                                                                                                                                                                                                             808 ------EBKEKEREREREKEREKEREKERELERGVKLAÇEGRAP-----VECPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SWECKER & MATHIS, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: Virginia
COUNTRY: United States
ZID: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
MEDIUM TYPE: FLOPPY disk
OFFRATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90/13101
APPLICATION NUMBER: FR 90/13101
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-0CT-1991
PRIOR APPLICATION NUMBER: WO 80/039,320
APPLICATION NUMBER: US 08/039,320
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
APPLICATION NUMBER: US 08/039,576
APPLICATION NUMBER: US 08/403,576
                                                                                                                                                                                                                                                ---LEGSKLAKKRADLVEKVRREAEQRAR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017753-025
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/479,537A FILING DATE: 07-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 5, Application US/08479537A ; Patent No. 5861381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (703) 836-6620
(703) 836-2021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1084 LLRPHVFMPEVTPDMDYLP 1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: BURNS, DOANE, STREET: P.O. BOX 1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Teskin, Robin L. REGISTRATION NUMBER: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Re-
CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CITY: Alexandria STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
```

```
APPLICANT: VAN HERDE, GEORGE V.
APPLICANT: VAN HERDE, GEORGE V.
APPLICANT: VAN HIN, ALEXIS C.
APPLICANT: BOUNSTRA, JAN B.
APPLICANT: DE WOLF, FREDERIK A.
APPLICANT: MOOBROEK, ANDREAS.
APPLICANT: WIND, RICHELE D.
APPLICANT: WIND, RICHELE D.
APPLICANT: WIND, RICHELE D.
APPLICANT: WIND, RICHELE D.
TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE FILE OF INVENTION: PREPARATION THEREOF
FILE REFERENCE: 2728-2
CURRENT APPLICATION NUMBER: US/09/219,849
NUMBER OF SEQ ID NOS: 50
SOFTHARE: PALENTIN V. 50
                       197 YP----PTYKAKPSYPPTYKAKPSYPPTYKAKPSYPSLIKAKPSYPPTYK 252
                                                                                       253 AKPTYKAKPTYKAKPTYPSTYRAKPTYP----PTYKAKPSYPPTYKAKPTYKAKPTYKAK 308
                                                                                                                                                        SAPTTIKEPA----PITIKSAPITPKEPSPITIKEPA--PITPKEPA----PITPKK 551
                                                                                                                                    PA--PTTPKEPA--PTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEK 607
                                                                                                                                                                                                      LAPTTPE--KPAPTTPEELAPTTPEEPT--PTTPEEPA-PTTPKAAAPNTPKEPAPTTPK 662
                                                                                                                                                                                                                           EPA-PIT----PKEPAPITPKETAPIT--PKGTAPITLKEPAPITPKKPAPKELAPITIK 715
                                                                                                                                                                                                                                                                                                                                                              716 EPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA--PTTPKKP 773
                                                                                                                                                                                                                                                                                                                                                                                                            774 APKELAPITIKGPISITSDKPAPITPKEJAPTIPKEPA--PITPKKPA-PIT----PETP 826
                                                                                                                                                                                                                                                                                                                                                                                                                                     577 PTYKAKPSYPPT-----YKAKPSYPPTYKAKPTYKA-----KPTYPST---YKAKPS 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               827 PPTTSEVSTPTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPE 886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; OTHER INFORMATION: Description of Artificial Sequence: Illustrative Y. OTHER INFORMATION: amino acid sequence US-09-219-849-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6.1%; Score 452; DB 4; Length 960;
25.6%; Pred. No. 1.9e-20;
ative 60; Mismatches 309; Indels 134;
TPKEPAPTAPKKPA--PTTPKEPA--PTTPKEPAPTTKEPSPTTP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    887 MITTAKDKTIERDLRTTPETTTAAP 911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   621 УРРТНКАКРТУКАКРЅУРРТНКАКР 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 25.6'
Matches 173; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 5
LENGTH: 960
                                                                 505
                                                                                                                                    552
                                                                                                                                                                   309
                                                                                                                                                                                                      608
                              g
                                                                 ð
                                                                                              qq
                                                                                                                                  ö
                                                                                                                                                              g
                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                       임
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
127 GPPGSRDPGPPGAPGPPGSRDPGPRGPPGSRDPGSPGPPGSRDPG 186
                                 72 PGPAGPPGSRDPG--PPGAPGPAGP--PGSRDPGPPGAPG-PAGPPGSRDPGPPGAPGPA 126
PTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAP--TTTKS 364
                                                                                                                                                                                                   244
                                                                                                                                                                                                                                                                                                               562
                                                                                                                                                                                                                                                                                                                                                                           563 TTP---KEPAPTTTKKPA-----PTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTT 612
                                                                                                                                                                                                                                                                                                                                                                                                   349 AGPPGSRDPGAPGAPGAPGSRDPGAPGAPGBPGSRDPGP----PGAPGPPG 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                     613 PEKPAPT-TPEELAPTTPEEPTPTTPEEPAPTTPKAAA-PNTPKEPAPTTP---KEPAPT 667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 405 SRDPGPPGAPGPPGSRDPGP--PGAPGPPGSRDPGPPGAPGPPGAPGPPGSRDPGPP 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TPKEPA-----PTTPKET-APTTPKGT-----APTTLKEPAPTTPKKP-APKELA 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     581 GP--PGAPGPPGSRDPGPPGAPGPPGSRDPGPPGAPGPPGAPGPPGSRDPGP--PGA 636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           818 PAPTTP----ETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALEN----S 868
                                                                         APTTPKEPAPTTTKEPA-PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP---KKPA
                                                                                                                                                                         460 ------PTAPK-KPAPTTP---KEPAPTTPKEPA-PTTTKEPSPTTPKEPAPTTTKS
                                                                                                                                                                                                                                                                                                 506 APTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTP---KKPAPTTPKEPAP
                                                                                                                                                 421 РТТРКЕРАРТТР---КЕРТРТТРКЕРА--РТТКЕРАРТТРКЕРА-------
                                                                                                                                                                                                                                                   245 PAGPKGAHGPAGPKGAPGPPAGPPGSPGAPGPAGPPGSRDPGP---A
                                                                                                                                                                                                                                                                                                                                 299 GPPGSRDPGP-----PGAPGPAGPPGSRDPGP--PGAPGPAGPPGSRDPGP-----PGAPGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   463 GAHGPAGPKGAHGPAGPKGAHGPAGPKGAHGPAGPKGAPGPAGPPGSRDPGPPGAPGPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTTTKEPTSTTSDKPAPTTPKGTA-PTTPKEPAPTTP---KEPAPTTPKGTA-PTTLKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           523 PPGSRDPGPPGA--PGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPPGSRDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 766 APTTPKKPAP-----KELAPTTTKGPTST--TSDKPAPTTPKETAPTTPKEPAPTTPKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Bredesen, Dale E.
APPLICANT: Rabizadeh, Sharroz
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
TITLE OF INVENTION: Polypeptides and Methods of Use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, V
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 23, Application US/09041886; Patent No. 6235872; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             869 РКЕРGVРТТКТРААТК 884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              California
: United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: United State
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: San Diego
STATE: Californ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-041-886-23
                                                                                                                                                                                   187
307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           711
                               QQ
                                                                       δ
                                                                                                   g
                                                                                                                                               ò
                                                                                                                                                                               g
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
```

41;

Gaps

564

g

g δy

g ò g ò

```
SYPPTYKSKPTYKPKITYPPTYKAKPSYPPTYKPKKTYPPT-YKPKLTYPPTYKPKASYP 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  286 TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTPKE----PASTTPKE 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKITTAKPINPRPSLPPNSDISKETSLIVNKETIVETKETTITNKQTSTDGKEKTISAKE 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             650 LPIIPQKGNNIPSNLPENPSDSEVEYPRPNDNGENSNNTMKSKKNI---PNEPIPSPGDN 706
                                                                                                                                                                                                                                                                                  627 ITPEEPTPTTPEEPAPTTPKAAAPNTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 685
510 TKEPAPTITKSAPTIPKEPSPITIKEPAPTIPKEPA-PTIPKKP-APTIPKEPAPTIPKE 567
                                                                     PTPTIKSAPT-TPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTTKSAPTTPKEPAPTTPKKPAPT-TPKEPAPTT----PKEPTPTTPKEPAPTTKEPAPT
                                                                                                                                                                                                                                                            686 KGTAPTILKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT
                                                                                                                                                                                                                                                                                                                                                                   ----NPEESNPKEP----INPEDNENPLIIQDEPIEPRNDSNVIPI
                                                                                                                                                                                                                                                                                                                                                                                                    806 TPKEP----APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                                     PAPTTTKKPAPTAPKEP - APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEBLAP
                                                                                                                                                                                                            524 SNPNE--PLNPNEP------SNPNEPSNPNEPSNPREE--PSNPKE--PSNPNE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 169;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 652;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: MAUGH, KATHY J.:ANDERSON, DAVID M.; STRAUSBERG, SUSAN L.: MCCANDLISS, RUSS:WEL, TENA; FILPULA, DAVID TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6.4%; Score 476.5; DB 6;
29.7%; Pred. No. 3.7e-22;
Live 71; Mismatches 284;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PYKGHEERIPKPHRSNDDYVYDNNVNKNNKDEPEIPNNE 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --EPGVPTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/528,762
FILING DATE: 25-MAY-11990
APPLICATION NUMBER: 82,456
FILING DATE: 07-AUG-1987
APPLICATION NUMBER: 933,945
FILING DATE: 24-NOV-1986
APPLICATION NUMBER: 650,128
FILING DATE: 13-SEP-1984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES: 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                   607 PKEPI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 860 PTPKALENSPK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Patent No. 5202236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; SEQ ID NO:13:
; LENGTH: 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Si
Matches 221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5202236-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5202236-13
                                                                                                                                                                   493
                                                                                                                                                                                                                                                                                                            265
                                                                                                                                                                                                                                                                                                                                             746
                                                                                                                                  568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                               δ
                                                            Qy
                                                                                             pp
                                                                                                                                  QY
                                                                                                                                                                   q
                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            343 TPTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTTKE 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   273 TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                      IPKEP-----APTIPKKPAPTIPETPPPTISEVSTPTTIKEPTTIHKSPDESTPELSAE 859
                                                                                                                                                                              LPIIPQKGNNIPSNLPENPSDSEVEYPRPNDNGENSNNTWKSKKNI---PNEPIPSPGDN 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAPTTTKSAPTTPKEP-APTTPKKP-APTTPKEP-APTTPKEP-TPTTPKEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          452 PITPKEPAPTAPK--KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sporozoite surface protein encoding
                                                                                  PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT
                686 KGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT
                                    PKEPI------NPEESNPKEP---INPEDNENPLIQUEPIEPRNDSNVIPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 826;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.6%; Score 488.5; DB 5;
26.4%; Pred. No. 8.9e-23;
iive 54; Mismatches 223;
                                                                                                                                                                                                                              ----EPGVPTTK 878
                                                                                                                                                                                                                                                    STREET: A. David Spevack
STREET: NMRDC Building 1 T-12 National Naval
STREET: Medical Center
CITY: Bethesda
STATE: Mr
                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Charoenvit, Yugin
APPLICANT: Charoenvit, Yugin
APPLICANT: Hedstrom, Richard
APPLICANT: Hedstrom, Richard
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria spc
TITLE OF INVENTION: Immunogen and gene en
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/00018
FILING DATE: 19920103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: PatentIn Release #1.24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                          RESULT 8
PCT-US92-00018-2
PCT-US92-00018-2
Sequence 2, Application PC/TUS9200018
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASAILALCATION: 4.24
ATTORNEY AGENT INFORMATION:
NAME: Spevack, Avram D.
TELECOMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 antino acids
TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 6.6%
Best Local Similarity 26.4%
Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             , MOLECULE TYPE: protein PCT-US92-00018-2
                                                                                                                                                                                                                                      PTPKALENSPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ZIP: 20814-5044
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE: 19
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                       707
                                                                                             746
                                                                                                                                                                  908
                                                                                                                                                                                                    650
                                                                                                                                                                                                                                      860
                                                                                                                                209
                                                            565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
```

45;

```
TTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEP 435
                                                                  495
                                                                              332 ที่รับราหารกราชการการกระการการกระการกระการกระการกระการกระการกระการกระการกระการกระการกระการกระการ
                   ---PSPTT-----TKE 535
                                                                                                                          ----- 527
                                                                                                                                                                                               438 VIGKDGRIENGMAFTMIPNDDTHVRFRFKVKDVGNTISVRCRKGAGKLEFPDRSLDFTIP 497
                                                                                                                                                                                                                                                                                                                                                     9/9
                                                                                                                                                                                                                                                                                                                                                                                 959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  607
                                                                                                                                                                                                                                                                                                                                                                                                                                        711
                                                                                                                                                                                                                                                                                                                                                                                                                                                       --TISDKPAPTIPKG-TAPTIPKEPAPITPKEPAPTIPKGTAPTILKE---PAPTIPKKP 773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LLDSNDEPIPGSQAGQIADISNLFPVQTHKSTGLPIDPMVGLPFDPKSGNLVHPYTNQTM 771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      995 INPTNNNTMDSSFAGAYKYAVSNGIKTDNVYGLPVGEITGLPKDPGSDIPFNSTTGELVD 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1055 PSTGKPINNSTAGIVSGKPGLPPIEDENGNLFDPSTNLPIDGNNQLVNPFINSTVSGSTS 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --INP--ETGKVIPGSLPGSLNYPSFNTPQQTDEITGKPVDTVTGLPYDPSTGEIIDPAT 1228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            774 APKELAPITIKGPISITSDKPAPITPKETAPITPKEPAPITPKKPAPITPETPPPITSEV 833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           944 TTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKP--EETAKPKD------RATNSKATT 994
                                                    436 ТРТТРКЕРАРТТКЕРАРТТРКЕРАРТАРККРАРТТРКЕРАРТТРКЕРАРТТТКЕРЅРТТР
                                                                                                                                                                                                                                                 PVAGHNSCSIIVGVSGDGKIHVSPYGSKDVSLIS----APIQPSELFNEVYCDTCTAKYG
                                                                                                                                                                                                                                                                                  ---Paptipkepaptitkkpaptapkepaptipketaptipkkliptipeklaptipekp
                                                                                                                                                                                                                                                                                                                                       APTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                            554 AIHSGYOTSADFVTTTTAKPTTTT--TGAPGOPTTTTTGSPSKPTTTTTKATTTT----
                                                                                                                                                                                                                                                                                                                                                                    --TTLNPIITTIOKPTTTT-----TT----KVPGKPPIATTTTIKPIVTTTTKATTT
                                                                                                                                                                                                                                                                                                                                                                                                                         657 TTTTVPTT----TTTTKRDEMTTTTTPLPDIGDIEITPIPIEKMLDKYTRMIYDYNSGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SCLSVSYLAAKNLTVDTDETYG--LPIDTLTGYPLDPVSLIPFN--PETGELFDPISDEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               828 MNGTIAGIVSGISASESLLSQKSALIDPATNMVVGEFGGLLNPATGVMIPGFLGPSEQTQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       888 FSPEIEDGGIIPPEVAAANADKFKLSIP--PSVP-----ESIPE-----KDQKIDSISE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------DLRTTPETTTAAPKMIKETATTTEKTTESKITATT-TQVTSTTTQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          935 LAYDIESGRLIGQVSRRPIPGSIAGDLNPIMKTPTQTDSVTGKPIDPTTGLPFNPPTGHL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1064 VEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPD-----MDYLPRVPNQGIIINP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KLPIPGSVAGDEILTEVLNITTDEVTGLPIDLETGLPRDPVSGLPQLPNGTLV 1281
                                                                                                                                                                                                                                                                                                                                                                                               PKETAPTTPKGTAPTTLKEPAPTTPKKPAPK----ELAPTTKEPTS-----
                                                                                                                                                                                                                          P------PREPAPTTPKKPAPTTPKE-------
                                                                                                           496 KEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STPT-----TTKEPTTIHKS------
                                                                                                                                                                                                                                                       498
                                                                                                                                                                                                                                                                                    560
                                                                                                                                                                                                                                                                                                                                          617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        772
                                                                                                                                                                                                                                                                                                                                                                     608
                                                                                                                                                                                                                                                                                                                                                                                               677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1229
ò
                         g
                                                                             g
                                                      à
                                                                                                            ö
                                                                                                                                      g
                                                                                                                                                                                            윱
                                                                                                                                                                   ô
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

```
34;
Sequence 2, Application US/07638431

Sequence 2, Application US/07638431

Sequence 2, Application US/07638431

SEMENAL INFORMATION:

APPLICANT: Hoffman, Stephen L.

APPLICANT: Hedstrom, Richard

APPLICANT: Khusmith, Srisin

APPLICANT: Khusmith, Srisin

APPLICANT: Mogers IV, William O.

TITLE OF INVENTION: Protective malaria sporozoite surface protein

TITLE OF INVENTION: Immunogen and gene

TITLE OF INVENTION: Immunogen and gene

CORRESPONDENCE ADDRESS:

ADDRESSEE: A. David Spevack

STREET: NAMPO Building I T-12 National Naval

STREET: Medical Center

CITY: Bethesda

STATE: MD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               343 TPTTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTKE 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         396 PAPTITKSAPITPKEP-APITPKKP-APTIPKEP-APTIPKEP-IPTIFKEPAPITKEPA 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             568 PAPITIKKPAPIAPKEP-APIIPKETAPITPKKLIPTIPEKLAPITPEKPAPITEELAP 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 627 TIPEEPTPTTPEEPAPTTPKAAAPNTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTTPKEPAPTAPK--KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTKSAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              510 TKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA-PTTPKKP-APTTPKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          493 PS------NPNEPSNPNEPSNPNE--PSNPKK-------PSNPNE--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 826;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
6.6%; Score 488.5; DB 1
Best Local Similarity 26.4%; Pred. No. 8.9e-23;
Matches 153; Conservative 54; Mismatches 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: US/07/638,431
FILING DATE: 19910110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWARE: Patentin Release #1.24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME: Spevack, Avrom D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: SDEVACK, AUTOM D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       826 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                              ZIP: 20814-5044
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; MOLECULE TYPE: protein US-07-638-431-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 19
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
```

```
PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS, THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM SPECIES INFECTIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         316 KGPALITPKEPTPTFPKEPASTTPKEPTPTTKSAPTTPKEPAPTTTKSAPTTPKEPAPT 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 DPYSNCPEN-----PUTGNLVSRSTGKTIPN----TYAGVYRSN-ETKTTEPSANT-- 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : | | : | : | : | : | 189
130 YAGVYESNETKTTEPSANTNFLLVDPKINAPCNSENSFEGGQIFDMGSKVYIPYTKCVGV 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSTIWKIKS----- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161 --SKNSAANRELQKKLKVKDNKKNRTKKKPTPRPPVVDEAGSGLDNGDFKVTTPDTSTTO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           256 KETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SITSGELKDPNKQAT-----ISGSRSCGWKQGYSIDSSTGFRVDSITGLPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1721;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 3;
                                                                                                                                                                                                                                                                                                                                                                                                            Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  219 HNKVSTSPKITTAKP-----INPRPSLPPNSDTSKE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.7%; Score 498.5; DE 21.8%; Pred. No. 5e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             480.76-1(HV)
                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION NUMBER: US/AP/AP/CURENT APPLICATION NUMBER: US/AP/AP/CETING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                    JMBER: US/08/928,361B
12-SEP-1997
                                                                                                                                                                                                                      STREET: 385 Sherman Avenue, Suite & STATE: CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---APATTPEDDDIDP--PLPNDPINP 699
                                                                                 US/08928361B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Verny, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1721 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 21.8
Matches 300; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             650-324-1678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein
                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                GENERAL INFORMATION:
APPLICANT: Petersen,
                                                                                                                                             TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                   Sequence 6, Application patent No. 6071518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                  COUNTRY: USA
ZIP: 94306-1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-928-361B-6
                                                   RESULT 6
US-08-928-351B-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38
      678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
      pp
                                                                                                                                                                                                                                                                                                                                                                                                                      32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTAPITPKEPAPITPKEPAPITPKGTAPTILKEPAPITPKKPAPKELAPITTKGPISITS 791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADFSIDHLLPHESDCGQYLQCVHGQTIARPCPGNLHFSPATQSCESPVTAGCQVFECDSD 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPK 731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302 DGEISPAPPVTEGNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERR 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----KSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEP 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      РАРТТРКЕРТРТТРКЕРАРТТКЕРАРТТРКЕРАРТАРККРАРТТРКЕРАРТТ - РКЕРА - - 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TIPKEPAPTITKEPAPTIPKEPAPTITKEPAPTITKSAPTIPKEPAPTIPKKPAPTIPKE 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          307 PTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTTIKSAPTTPKEPAPTTKSAP 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     557 ---PKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNT--PKEPAPTTPKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAPTAAPTAAPTTAVPEIPTTVTSPPTAAPTTAAP---APNTT------VTVPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELLPNGCPADFDIHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                362 CGAGTHFSFELQQCDHIELVGCTLPGGESEEVDVDEDA-CTGWYCPTEPIEWEPLPNGCP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --PTTTKEPSPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----APTTPKKPAPTT-------
                                                                                                                                                                                                                                                                                                                                                                                       Length 786;
                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                         6.7%; Score 498.5; DB 4; nilarity 29.1%; Pred. No. 2e-23; Conservative 25; Mismatches 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                633 TAAPIT-AAPAPNTIVTVPPTAAPTAAPPTVAH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DKPAPTTPKETAPTTPKEPAPTTPKKP 818
                                              NAME: Michaels, Christopher A REGISTRATION NUMBER: 34,390 REGISTRATION NUMBER: 34,390 TELECOMMUNICATION INFORMATION: TELEPHONE: (607) 256-2000 TELEPRAX: (607) 256-3628 INFORMATION FOR SEQ ID NO: 3: SEQUENCE CHARRATERISTICS: LENGTH: 786 amino acids
                                                                                                                                                                                                                                                                                                             ORGANISM: Trichoplusia ni
TISSUE TYPE: peritrophic membrane
                      ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                              protein
NO
                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         537 APTT----PKEP--
                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                               linear
                   CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----EPAPTT-
                                                                                                                                                                                                                                                  MOLECULE TYPE:
                                                                                                                                                                                                                                                                                   ON
N
                                                                                                                                                                                                                                                                                                 ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                   HYPOTHETICAL:
                                                                                                                                                                                                                                      TOPOLOGY:
                                                                                                                                                                                                                                                                                      ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                 US-09-103-429A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                  200;
                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Si
Matches 200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  367
```

δ g δy qq δ

qq

Qγ

g

g Ω

δ qq δλ g οy g δý g

οy

827 853 887 868 934

```
995 INPTNNNTMDSSFAGAYKYAVSNGIKTDNVYGLPVGEITGLPKDPGSDIPFNSTTGELVD 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1115 GTTKPKPGIPVNGGGVVPDEEAKDQADKGKDGLIVPPTNSINKDPVTNTQYSNTTGNI-- 1172
                                                                                      PSTGKPINNSTAGIVSGKPCLPPIEDENGNLFDPSTNLPIDGNNQLVNPETNSTVSGSTS 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1064 VEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPD------MDYLPRVPNQGIINP-- 1113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --TISDKPAPTTPKG-TAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKE---PAPTTPKKP 773
                                                                                                                                                             712 LLDSNDEPIPGSQAGQIADTSNLFPVQTHKSTGLPIDPMVGLPFDPKSGNLVHPYTNQTM 771
                                                                                                                                                                                             774 APKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEV 833
                                                                                                                                                                                                                                                                                                                                                                                                  935 LMYDIESGRLIGQVSKRPIPGSIAGDLNPIMKTPTQTDSVTGKPIDPTTGLPFNPPTGHL 994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   944 TTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKP--EETAKPKD------RATNSKATT 994
                                                                 PKETAPTTPKGTAPTTLKEPAPTTPKKPAPK----ELAPTTTKEPTS-------
                                                                                                                                                                                                                        Sequence 3, Application US/09103429A
Patent No. 6187558
GENERAL INFORMATION:
APPLICANT: Granados, Robert R
APPLICANT: Wang, Ping
TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                              -LSECT---------
                                                                                                                                                                                                                                                                                              828 MNGTIAGIVSGISASESLLSQKSALIDPATNMVVGEFGGLLNPATGVMIPGFLGPSEQTO
                                                                                                                                                                                                                                                                                                                                -----PELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTAKDKTTER---
                                                                                                                                                                                                                                                                                                                                                              888 FSPEIEDGGIIPPEVAAANADKFKLSIP--PSVP-----ESIPE-----KDQKIDSISE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1114 -----PVDGLTTLRNGTLV 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1229 KLPIPGSVAGDEILTEVLNITTDEVTGLPIDLETGLPRDPVSGLPQLPNGTLV 1281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Brown, Pinnis1 & Michaels, P.C. STREET: 118 No. 6187558th Tioga CITY: Ithaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 14850
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Vers
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,429A
FILING DATE: 24-JUN-1998
                                                                                                                                                                                                                                                                834 STPT----TTKEPTTIHKS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-103-429A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CITY:
STATE:
                                                                                                                               720
                                                                                                                                                                                                                                                                                                                                                                                                   899
                                                                                                                                                                                                                                                                                                                                  854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1055
                               αq
                                                               ŏ
                                                                                                                        οý
                                                                                             q
                                                                                                                                                             g
                                                                                                                                                                                                 à
                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                           DNAS AND RNAS
                                                      APPLICANT: LEECH, JAMES
APPLICANT: LEECH, JAMES
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: NELSON, RICHARD, C.
APPLICANT: OUT, JIR
TITLE OF INVENTION: VOR PROPHYLAXIS AND TREATMENT OF CLYPCOPROTEINS, DNAS AND TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF CLYPCOPROTEINS, DNAS AND TITLE OF INVENTION: NOWBER: US/08/700,651B
CURRENT FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751
EARLIER PILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PALENTIN VOR: 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              107 STIKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKS----- 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --SKNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQ 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DPYSNCPFN-----PVTGNLVSRSTGKTIPN----TYAGVYRSN-ETKTTEPSANT-- 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HNKVSTSPKITTAKP-----INPRPSLPPNSDTSKE-------TSLTVN 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    256 KETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTT 315
                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             316 KGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTKSAPTTPKEPAPT 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     436 TPTTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTP 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     496 KEPAPITIKSAPITIKEPAPITIKSAPITIPKE---------------527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   438 VIGKDGRIENGMAFTMIPNDDTHVRFRFKVKDVGNTISVRCRKGAGKLEFPDRSLDFTIP 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTT 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              38 SITSGERKDPNKQAT------ISGSRSCGWKQGYSIDSSTGFRVDSITGLPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             376 TTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTFKEPAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             212 รากกราชการกรรกราชการกรรรกรรรรรกรราชการกรรรกราชการกรรรรร
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 536 Р---------АРТТРКЕРАРТТРККРАРТТРКЕ-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PVAGHNSCSIIVGVSGDGKIHVSPYGSKDVSLIS----APIOPSELFNEVYCDTCTAKYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               554 AIHSGYQTSADFVTTTTAKPTTTT--TGAPGQPTTTTTGSPSKPTTTTTKATTTT----
                                                                                                                                                                                                                                                                                                                                                                                                   6.7%; Score 499.5; DB 3; Length 1721;
21.8%; Pred. No. 4.4e-23;
tive 94; Mismatches 510; Indels 469;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190 KHTTTTTTTTTTTT--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Application US/08700651B 6015882
                                                                                                                                                                                                                                                                                                             ; TYPE: PRT; CRYPtosporidium parvum US-08-700-651-5
          Patent No. 6015882
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
                                                                                                                                                                                                                                                                                                                                                                                            Query Match 6.7%
Best Local Similarity 21.8%
Matches 300; Conservative
                                                                                                                                                                                                                                                                              SEQ ID NO 5
LENGTH: 1721
Sequence 5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
```

ò g

g ò

ð 8 õ g

```
1183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------NSKATTPKFQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTS 1033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1184 GIVSGKRGLPPIEDENGNLFDPSTKLPIDGNNQLVNPETNSTVSGSTSGSTKPKPGIPVN 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1298 IPGSLPGSLNYPSENTPQQTDEITGKPVDTVTGLPYDPSTGEIIDPATKLPIPGSVAGDE 1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1025 DKFKLSIPPSVPESIPEKDQKIDSISELMYDIESGRLIGQVSKRPIPGSIAGDLN----P 1080
                                                                                                                                                                                                                                                                                                                                                                                                                 -----LSAEPTPKALENSPKEPGVP 875
                                                                                                                                                                                                                                                                                                                                                                 908 TYGLPIDTLTGYPLDPVSLIPFNPETGELFDPISDEIMNGTIAGIVSGISASESLLSQKS 967
                                                                                                                                                                                                                                                                                                                                                                                               ----TAPTTPKEPAPTTPKKPAPTTPETP-----PPTTSEVSTPTT 838
                                                                                                                                                                                                                   696 PAPTTPKKPAPK----ELAPTTTKEPTS------TTSDKPAPTTPKG-TAP 735
                                                                                                                                                                                                                                      198 DEMITITIPLEDIGDIEITPIPIEKMLDKYTRMIYDYNSGLLLDSNDEPIPGSQAGQIAD 847
                                                                                                                                                                                                                                                                                                                                     -----KGPTSTTSDKP 794
                                                                                                                      PITTI--TGAPGOPTTITGSPSKPTTTTTKATTTT----TILNPIITTTOKPTTT 741
                                                                                                                                                                              636 TPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKE 695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1076 AEGETPHMLLRPHVFMPEVTPD------MDYLPRVPNQGIIINP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AYKYAVSNGIKTDNVYGLPVDEITGL-PKDPVSDIPFNSTTGELVD--PSTGKPINNYTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            876 TIKTPAATKPEMTTAKDKTTERDLRTTPETTTAAPKMTKETA----TTTEKTTESKITA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTTQVTSTTTQDTTPFKITTLKTTTLAPK - - VTTTKKTITTTEIMNKPEETAKPKDRAT -
                                                                                                    848 ISNLFPVQTHKSTGLPIDPMVGLPFDPKSGNLVHPYTNQTMSGLSVSYLAAKNLTVDTDE
                570 MIPNDDTHVRFRFKVKDVGNTISVRCGKGAGKLEFPDRSLDFTIPPVAGHNSCSIIVGVS
                                              -----PAPTTPKEPAPTTKK
                                                                        630 GGGKIHVSPYGSKDVSLISAPIQPCELFNEVYCDTCTAKYGAIHSGYQTSADFVTTTAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG, SUSAN D.; MCCANDLISS, RUSS; WEI, TENA; FILPULA, DAVID TILLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1114 MLSDETNICNGK------PVDGLTTLRNGTLV 1139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1358 ILTEVLNITTDEVIGLPIDLETGLPRDPVSGLPQLPNGTLV 1398
                                                                                                                                                                                                                                                                                                                                        --GTAPTILK----EPAPTTPKKPAPKELAPTTT------
                                                    ----PAPTTPKE-----
                                                                                                                                                                                                                                                                                    TTPKEPA-----PTTPKEPAPTTPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES: 39
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/528,762
FILING DATE: 25-MAY-1990
APPLICATION NUMBER: 82,456
                                                                                                                                                                                                                                                                                                                                                                                                                                                               839 TKEPTTIHKSPDESTPE------
                                                                                                                                                                                                                                                                                                                                                                                                         795 APTTPKE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5202236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           520238-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   932
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1034
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           989
                                                                                                                                                                                                                                                                                         736
                                                                                                                                                                                                                                                                                                                                                  756
                                                                                                                                                                                                    742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; PROTEIN
                                                                                                                216
                                                                                                                                            900
                                                       552
552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                                                         Q\underline{\gamma}
                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                         g
                                                                                                                                                                        δ
                                                                                                                                                                                                  g
                                                                                   g
                                                                                                                δŽ
                                                         δλ
```

```
50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 654 KEPAPTTPKEPAPTT----PKEPAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAPK 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        323 SYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKVKPTYKAKPTYPS 382
                                                                                                                                                                                                                                                                                                                                                                    -APTIPKEPA--PIAPKKPA--PITPKEPA--PITPKEPA--PITTKEPS--PITPKEPA 499
                                                                                                                                                                                                                                                                                                                                                                                                                   -PTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPA---PTTPKEPA----PTTPKKP 552
                                                                                                                                                                                                                                                                                                                                                                                                                                              322
                                                                                                                                                                                                                                                                                  93 SYPPTYKSKPITYRPKITYPPTYKAKPSYPPTYKPKKTYPPT-YKPKLTYPPTYKPKASYP 151
                                                                                                                                                                                                                  286 TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKE----PASTTPKE 341
                                                                                                                                                                   226 PKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTTSAKE 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     767 -PITPKKPAPKELAPITIKGPI--SITSDKPAPIIPKETAPTIPKEPAPIIPKKPA--PI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---TPETPPPTTSEVSTPTTKE----PTTIHKSPD-----ESTPELSAEPT-PKALENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            602 YKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    869 PKEPGVPTTKTPAA-----TKPEMTTTAKDKTTERDLRTTPETTTAAP-----KMT
                                                                                                                                                                                                                                                                                                                                                                                                                                   PEKLAPTTPE--KPAPTTPEELAPTTPEEPT---PTTPEEPA-PTTPKA--AAPNT----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ELAPTTIKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A--PITPKEPA--PITPKEPA--PTITKKPA--PTAPKEPAPTTPKETAPTTPKKLTPTT
                                                                                                                                                                                                                                    10 YPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPS
                                                                                                                                                                                    342 PTPTTIKSAPT-TPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPA--P
                                                                                                                                                                                                                                                                                                                     TTTKSAPTTPKEPAPTTPKKPAPT-TPKEPAPTT----PKEPTPTTPKEPAPTTKEP---
                                                                                                                                            Indels 175;
                                                                                                                       Length 744;
                                                                                                                       6.8%; Score 506.5; DB 6;
29.9%; Pred. No. 6e-24;
tive 78; Mismatches 308;
FILING DATE: 07-AUG-1987
APPLICATION NUMBER: 933,945
FILING DATE: 24-NOV-1986
APPLICATION NUMBER: 650,128
FILING DATE: 13-SEP-1984
SEQ ID NO:25:
LENGTH: 744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               720 YKAKPIYPSTYKAKPIYPST 739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           915 KETATTTEKTTESKITATTT 934
                                                                                                                                                   Conservative
                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
US-08-700-651-5
                                                                                                                                                     239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              605
                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                              399
                                                                                                                                                                                                                                                                                                                                                                               451
                                                                                                                                                                                                                                                                                                                                                                                                                               200
                                                                                                                                                                                                                                                      54
                                                                                          520238-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QY
                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                     Dp
                                                                                                                                                                                                                                                                                                                               QΥ
                                                                                                                                                                                                    q
                                                                                                                                                                                                                               δŏ
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                               QΥ
```

Conservative

226;

Matches

```
83 HADASTISSAYSAPFELDVSGV--PIEPNIRRMVDPVSLMLFDNSTGVMYDPN---INSI 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 IESE--EITEEHSVSENQESSSSSSSSSSTIWKIKSSKNSAANRELOKKLKVKDNKKN 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              242 PNSDTSKETSLTVNKET-----TVETKETTTTNKQTST-----DGK-----EKTTSAKE 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPT 345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            346 TIKSAPTIPKEPAPTITKSAPTTPKEPAPITTKEPAPTTPKEPAPTTTKEPAPTTTKSAP 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |:|: ||| || ; || || 5
510 KRPTTTTATTTTSETESVIKPDEWCWLEKNGECEAKGATYVGVIGKDGRIENGMAFT 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 196 GWKQ-----ITGLPTDPYPNCP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FNPVTGNLVSRSTGKTIPNTYAGVYRSNETKTTEPSANTNFLLVDPKINAPCNSENSFEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   183 RTKKKPTPKPPVVDEAGSGLDNG-DFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 513; DB 3; Length 1837;
Pred. No. 6.8e-24;
5; Mismatches 539; Indels 470;
                                                                                                                                        SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928.361B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HNPTSPPSSKKAPP----PSGASQTIKSTTKRSPKP----
                                                                                                                                                                                                                                                                                                                        480.76-1(HV)
                                                                                                                                                                FILING DATE: 12-SEP-1997
CLASSIFICATION:
                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTONNEY AGENT INFORMATION:
                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      526 КЕРЅРТТТКЕРАРТТРКЕРАРТТРКК----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 21.6%; Pre
Matches 307; Conservative 105;
                                                                                                                                                                                                                                                                                   NAME: Verny, Hana
REGISTRATION NUMBER: 30,518
                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: 480
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1678
TELEFAX: 650-324-1678
                                                                                                                                                                                                                                                                                                                                                             TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                     1837 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY: linear
MOLECULE TYPE: protein
                                                            94306-1840
                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid STRANDEDNESS:
                  Palo Alto
                                                USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-928-361B-5
                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 5, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
TITLE OF INVENTION: SPECIES INFECTIONS
TITLE OF INVENTION: SPECIES INFECTIONS
TORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
           40;
                                                                                                               67 RDCAPGTEFKFSAQTCVHAALAGCTLPGPPAETT---QAPATTQ--APTTTQ--APTTTT 119
                                                                                           ----TKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTK 402
                                                                                                                                                      SAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTA 462
                           300 ISKVLAKPIPKAETTTKGPALTIPKEPTPITPKEPASTIPKEPTPTIKSAP-----TTP 354
                                                                                                                                                                                  120 QAPTITIQATITI ----- QAPTITIQAPITI ----QATITIQAPITIT ---- 156
                                                                                                                                                                                                                                       PKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAP 522
                                                                                                                                                                                                                                                                                                     254 IHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVECDGEISPAPPVTE 313
                                                                                                                                                                                                                                                                                                                                                                                                                           577
                                                                                                                                                                                                                                                                                                                                                                                                                                           314 GNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERRCGAGTHFSFELQ 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       374 OCDHIELVGCTLPGGESEEVDVDEDACTGWYCPTEPIEWEPLPNGCPADFSIDHLLPHES 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------PEKLAPITPEKPAP 618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      434 DCGOYLOCVHGOTIARPCPGNLHFSPATOSCESPVTAGCOVFECDSDNOCTSTAAPTAAP 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   494 TAAPTAAPTAAPTAAPSTVVPPA-TPPATAAPVPPTTAIPT----PAPTAAPTAAPTTAA 548
           Gaps
                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TIPEELAPITPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPK 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 736 TTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTKGPTST-TSDKP 794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             854
                                                       9 TALGLVAARPEVSDAEKNPALHEPHPDXP--PAEQXXLLPXEYDCTKFYYCEYGLKFIAP
                                                                                                                                                                                                                                                                                   TTPKEPSPTTTKEPAPTTP--KEPAPTTPKKPAPTTPKEPAPTTPK-------
                                                                                                                                                                                                                                                                                                                                                          --EPAPTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------TPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ETAPTT---PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               795 APTTEKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 342;
33; Mismatches 211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            855 ELSAEPTPKALENSPKEPGVPTTKTPAATKPE 886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       686 NTTAAP-----VTTTSAPATT-PE 703
                                                                                                                                                                                                                                                                                                                                                                                                              575 K-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-08-928-361B-5
                                                                                         355
                                                                                                                                                                                                                    463
                                                                                                                                                                                                                                                                                   523
                                                                                                                                                                                                                                                   157
                                                                                                                                                                                                                                                                                                                 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 599
                           ò
                                                                                       à
                                                                                                                    a
                                                                                                                                                                                 g
                                                                                                                                                    õ
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

49;

Sequence 3

```
TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ANTI-SENSE: NO
ORIGINAL SOURCE:
                                                                                                                                                                                                            RESULT 1
US-09-103-429A-4
                                                                                                371
370.5
370.5
370.5
370.5
370.5
404.5
404.5
398.5
398.5
383.5
383.5
5, Appli
23, Appli
5, Appli
2, Appli
2, Appli
2, Appli
13, Appli
13, Appli
53, Appli
68, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                    5, Appli
5202236
5, Appli
3, Appli
5, Appli
2, Appli
2, Appli
5, Appli
5, Appli
                                                                          (without alignments)
623.836 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
                                                                                                                 7410
1 QDLSSCAGRCGEGYSRDATC.....ARAITTRSGGTLSKVWYNCP 1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2
                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
Sequence
Sequence
                                                                April 26, 2002, 16:16:35; Search time 49.78 Seconds
                                                                                                                                                                                                                                                                       Issued_Patents_AA:*
.: /cgn2_c/ptodata/2/iaa/5A_COMB.pep:*
.: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
.: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
.: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
.: /cgn2_6/ptodata/2/iaa/BCOMB.pep:*
.: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
.: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                       Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-477-509B-103
US-08-642-255-101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-783-774-2
PCT-US95-04611A-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-479-537A-2
US-09-083-116-2
US-08-276-967-2
US-08-642-255-132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-707-237A-75
US-08-482-085B-103
US-08-397-633A-50
                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-103-429A-3
US-09-103-429A-3
US-08-928-361B-6
US-07-638-431-2
PCT-US92-00018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-08-397-633A-53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-07-609-716-65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-08-175-155-68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-041-886-23
US-08-479-537A-5
US-09-083-116-5
                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-103-429A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-219-849-5
            GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                   212252 seqs, 22503292 residues
                                                                                                      US-09-556-246-1_COPY_25_1404
                                                                                                                                                                                                                                  Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                 protein search, using sw model
                                                                                                                                            BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                      Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                               805
1837
1724
1721
1721
826
826
965
965
111867
11867
11867
2035
2476
837
837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                           Title:
Perfect score:
                                                                                                                                             Scoring table:
                                                    OM protein
                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                            Database
                                                                                                                              Sednence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
                                                                        Run on:
```

```
Sequence 2, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 114, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 28, Appl
                                                                                   36, Appli
2, Appli
4, Appli
6, Appli
6, Appli
6, Appli
8, Appli
84, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Granados, Robert R
APPLICANT: Wang, Ping
APPLICANT: Wang, Ping
TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin
TITLE OF INVENTION: cDNA and Related Products and Methods
TITLE OF INVENTION: cDNA and Related Products and Methods
MUNBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Pinnisi & Michaels, P.C.
STREET: 118 No. 6187558th Tioga
CITY: Ithaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STATE: NA

COUNTRY: USA

ZIP: 14850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OMPUTER: IBM PC COMPATIBLE
OMPUTER: IBM PC COMPATIBLE
OMPUTER: IBM PC COMPATIBLE
OMPUTER: PREADING NOTES:
SOFTWARE: PREADING NOTES:
APPLICATION NUMBER: US/09/103,429A
FILING DATE: 24-JUN-1998
FILING DATE: 24-JUN-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MAGNAGES, Christopher A
REGISTRATION NUMBER: BTI-39
REGISTRATION NUMBER: BTI-39
REGISTRATION NUMBER: BTI-39
TELEFAN: (607) 256-2000
TELEFAN: (607) 256-2000
TELEFAN: (607) 256-3008
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHRRACTER/STICS:
LENGTH: 805-8010 ACCIDENT OF ACCID
US-08-475-411A-65
US-08-478-029A-65
US-08-478-029A-65
US-08-397-633A-36
US-08-397-633A-36
US-08-397-633A-26
US-08-95-1061A-6
US-08-95-1061A-6
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-97-462-8
US-08-98-98-25-62
US-08-98-98-26
US-08-98-98-26
US-08-98-98-26
US-08-98-98-26
US-08-98-98-26
US-08-98-98-98-26
US-08-98-98-98-26
US-08-98-98-98-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 4, Application US/09103429A
Patent No. 6187558
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ) ORGANISM: Trichoplusia ni
) TISSUE TYPE: peritrophic US-09-103-429A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.0%;
27.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: protein HYPOTHETICAL: NO
   408
682
682
682
11537
11231
11848
1948
960
1786
762
762
1064
11187
11187
```

```
35;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKGTAP--TTLKEPAPTIPKKPAPKELAPTITKGPTSTTSDKPA-PTTPKETAPTTPKEP 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APTTPKKPAPTTPETPPPTTSEV--STPTTTKEPTTIHKSPDE---STPELSAEPTPKAL 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                691 -PEKAKSP----VKEEAKSPEKAKSPVKEEAKSPEKAKSPV----KEEAKTPEKAKSPV- 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              638 EEPAPTTPKAAAPNTPKEPAPTTP----KEPAPTTPKEPAPTTPKETAPTTPKGTAPTTL 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   694 KEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTT 753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    401 TKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAP 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    461 TAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKS 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           632 VKSPEKÄKSPTKEEÄKSPEKAKSPEKEEAKSPEKAKSPVKAEAKS-PEKAKSPVKAEAKS 690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          521 APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTA 580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            581 PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPE--KPAPTTPEEL-APTTPEEPTPTTP 637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               341 EPTPTIIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTT 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         222 VSTSPKITTAKPINPRPSLPPNSDTSKETSLTVVNKETTVETKETTTINKQTSTDGKE-KT 280
                                                                                                                                                                                                                                                                                                                                                                                                                               438 VKSEEKIKVVE-----KSEKETVIVEE-----QTEETQVTEEVTEEEEKEAKE 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                           281 TSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPK 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   481 EEGKEEEGGEEEBEAEG-----GEBETKSP----PAEEAASPEKEAKSPVKE 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119 KKTKKVIESE--EITEEHSVS-ENQESSSSSSSSSSSTIWKIKSSKNSAANRELQKKLK 175
                                                                                                                                                                                                                                                                                                                  VK------DNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 221
                                                                                                                                                                                                                                                                                                                                                                      6.8%; Score 502; DB 1; Length 1020;
26.7%; Pred. No. 3.4e-13;
tive 81; Mismatches 356; Indels 136; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KSPLKADAKAPEKEIPKKEEVKSPVKEEEKPQEVKVKEPPKKAEEEKAPATPK-----
                         InterPro; IPR001664; IF.
Promise Proposals, filament; 1.
PROSITE; PS00226; IF; 1.
Intermediate filament; Coiled coil; Neurone; Phosphorylation.
DOMAIN 101 413 ROD.
DOMAIN 414 1020 TAIL.
                                                                                   101 4.3 ROD.

414 1020 TAIL.

101 132 COIL 1A.

133 145 LINKER 1.

245 266 LINKER 12.

267 288 COIL 2A.

293 413 COIL 2B.

201 2B.

202 202 LINKER 2.

203 413 COIL 2B.
                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 26.7%
Matches 209; Conservative
PIR; S00979; QFHUH. MIM; 162230; -.
                                                                                                                                                                                                   DOMAIN
SEQUENCE
                                                                                            DOMAIN
                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     811
                                                                                                                       DOMAIN
                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                               176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΥ
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
g
                                                                                                          g
                                                                           ò
                                                                                                                                            δ
                                                                                                                                                                         qq
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; X74812,
EMBL; X7481; CAA2.
EMBL; X74870; CAA52862.1,
PIR; S21054; SNA_POI_A.

A MIM; 180666; -187000684; RNA_POI_A.

DR InterPro; IPR0002879; RNA_POI_A.

DR Pfam; PF00623; RNA_POI_A.

DR PROSTIE: PS00115; RNA_POI_A.

DR PROSTIE: PS00115; RNA_POI_A.

Transferase; DNA-directed RNA polymerase; Transcription; Zinc.finger.

NNA-binding; Nuclear protein; Phosphorylation; Zinc.finger.

To CAL2-TYPE (POTEWIAL).

TO 1958 CARBOXYL-TERMINAL).

TO 1958 CARBOXYL-TERMINAL).

D -> Y (IN REF. 2).

"W; 6876FC25692A657E CRC64;

"W; 6876FC25692A657E CRC64;
                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration the European Bioinformatics in the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial or send an email to license@lsb-slb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           47;
                                            Mita K., Tsuji H., Morimyo M., Takahashi E., Nenoi M.,
Ichimura S., Yamauchi M., Hongo E., Hayashi A.;
"The human gene encoding the largest subunit of RNA polymerase II.";
Gene 159:285-286(1995).
-I. FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
                                                                                                                                                                                                      -- SUBCELLULAR LOCATION: NUCLEAR.
-- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
-- THE PHOSPHORYLATION ACTIVATES POLZ.
-- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, AND POLYMERASE II FOR THE MENNA PRECURSOR, AND POLYMERASE II FOR THE MENS.
-- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1507 PSPMGGISPAMTPWNQGATPAYGAWSPSVGSGMTPGAAGFSPSAASDASGFSPGYSPAWS 1566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPTTPKEPASTTPKE-PTPTIKSAPTTPKEPAPTTTKSAPTTPKE-----PAPTTTKE 379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       438 TIPKEPAPITKEPAPTIPKEPAPTAPKKPAPTIPKEPAPTIPKEPAPTITKEPSPTIFKE 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      498 PAPITIKSAPITIKEPAPITIKSAPITPKEPSPITIKEPAPITPKEPAPITPKKPAPITP 557
                                                                                                                                                           CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE
                                                                                                                                                                                          SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        380 PAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEP--APTTPKEPTP
                              MEDLINE-95347616; PubMed-7622068;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; X63564; CAA45125.1; -
EMBL; X74874; CAA52862.1; -
EMBL; X74873; CAA52862.1; JOINED.
EMBL; X74872; CAA52862.1; JOINED.
EMBL; X74870; CAA52862.1; JOINED.
EMBL; X74870; CAA52862.1; JOINED.
                                                                                                                             OF DNA INTO
SUBSTRATES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1567
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
RX MEDLINE—88328981; Pubbed=3138108;
RA Lees J.F., Shneidman P.S., Skuntz S.F., Carden M.J., Lazzarini R.A.;
Lees J.F., Shneidman P.S., Skuntz S.F., Carden M.J., Lazzarini R.A.;
The structure and organization of the human heavy neurofilament
are ubunit (NF-H) and the gene encoding it.";

EMBO J. 7:1947-1955(1988).

C. PURCTION: NEUROFELIAMENTS USGALLY CONTAIN THREE IF PROTEINS: L, M,
EMBO J. 7:1947-1955(1988).

C. PURCTION: NEUROFELIAMENTS USGALLY CONTAIN THREE IF PROTEINS: L, M,
EMBO J. 7:1947-1955(1988).

C. PURCTION: NEUROFELIAMENT SUALLER NF PROTEINS
C. NF-H AS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT
PHOSPHORYLATED ON A NUMBER OF PRESENTING IN THE FORMATION OF
C. PTM: THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
C. THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
C. THOUGHER.
C. OF AXONAL CALIBER.

C. THOUSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
C. OF THE LARGER NEUROFILAMENT POLYPEPTINGS (NF-M AND NF-H), THE
C. LEVELS OF PHOSPHORYLATION BRING ALTERED DEVELOPMENTALLY AND
C. LEVELS OF PHOSPHORYLATION BRING ALTERED DEVELOPMENTALLY AND
C. SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation—the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                          SY-SPISPSY-SPISPSY-SPISPNY-SPISPNYT-SPISP-SYSPISPNY-TITLE 1777 T- TITLE SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPSY-SYSPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISPNYT-SPISP
                                                                                                                                                                                     TPKETAPTTPKGTAPTTLKEPAPTTPK-KPAPKELAPTTTKEPTSTTSDKPAPTTPKGTA 734
558 КЕРАРТТРКЕРАРТТТККРАРТАРКЕРАРТТРКЕТАРТТРККLТРТТРЕКLAРТТРЕКРА
                                                                                                                                          PTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEP--APT
                                                                                                                                                                                                                                                                                                                                                                                                                                       PTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT-STTSDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              794 PAPTTPKETAPTTPKEP--APTTPKKPAPTTPETPPPTTSEVSTPTTKEPTTIHKSPDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-OCT-1989 (Rel. 12, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1020 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X15306; CAA33366.1; -. EMBL; X15307; CAA33366.1; JOINED. EMBL; X15308; CAA33366.1; JOINED. EMBL; X15309; CAA33366.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-OCT-1989 (Rel. 12, 01-OCT-1989 (Rel. 12, 20-AUG-2001 (Rel. 40,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1966 SDEE 1969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 852 STPE 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NEFH OR NFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NFH_HUMAN
P12036:
                                                                                                                                                                                                            1778
                                                                                                                                              618
                                                                                                                                                                                                                                                                                         919
                                                                                                                                                                                                                                                                                                                                                                  1829
                                                                                                                                                                                                                                                                                                                                                                                                                                       735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NFH_HUMAN
ID NFH_H
```

```
960 KVTTTKKTITTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPK----T 1015
                                                                                                                                                                                                                                                                                            LRITPETITAAPKMIKETATITEKITESKITATITQVTSTITQDTIPFKITILKITILAP 959
                                                                                                                                                                                                                                     KEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERD 899
                                                                                                                                                                    670 KEPAPTIPKETAPTIPKGTAPTILKEPAPTIPKKPAPKELAPTITKEPTSTTSDKPAPTT 729
                                                                                                                                                                                                730 PKG-----TAPIT-----PKEPAPTIPKEPAPTIPKGTAPTILKEPAPTTPKKPAPKELA 779
                                                                                                                                                                                                             329 TKGCCFDSSIPQTKWCFYTLSQVADCKVEPSQRVDCGFRGIT----ADQCRQKNCCFDSS 384
327 IPITPKEPASITPKEPTPITIKSAPITPKEPAPITIKSAPTIPKEPAPITIKEPAPITPK 386
                                     104 GKAPATAAAPVPTTAASKAPTTAAAATHSTAAAAPTTAASAAKSKERSTSSSSEEEHCH 163
                                                       --PITKEPAPI---TPKEPAPIAPKKPAPTIPK------EPAPTIPKEPAPTITKE 489
                                                                                                              550 KKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLA 609
         511 TITITKATITITSGECKME----PSKRADCGYPGITESQCRSKGCCFDSSIPQTKWCFYS
                                                                                   490 PSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTP
                                                                                                                       ------REDCGYSGITESQCR
                                                                                                                                                                                                                            PITIKGPISITSDKPAPITPKETAPITPKEPAPITPKKPAPITPETPPPITSEVSTPTIT
                                                                                                                                                                                                                                                                    PTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTP
                                                                ------TTTTKATTTPTTTTTTP---------
                            EPAPITIKEPAPITIKS-APITPKEPAPITPKKPAPITPKEPAPITPKEPAFITPKEPA-
                                                                                                                                                                                                                                                                                                                                             1016 MPRVRKPKTTPTPR 1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                               RPB1_HUMAN
                                                                                                                                                                                                                                                                          TTTT
                                                                                                                                                                                                                                 780
                                                                                                                                                                                                                                                             840
                                                                                                                                                                                                                                                                          437
                                                                                                                                                                                                                                                                                        006
                                                                                                                                                                                        290
                                 387
                                                            445
                                                                                                      219
                                                                                                                                 246
                                                                                                                                              610
                                                                                                                                                            271
                   25
                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                          Pp
                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                        QΩ
                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                 οp
                                                                                                                                                           QQ
                                                                                                                                                                          Qγ
                                                                                                                                                                                                     δ
                                                                         g
                                                                                        δŏ
                                                                                                                   δy
                                                                                                                                               \tilde{\Omega}
                                              g
                                                             ò
                  9
                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                     1-8.
P-TYPE 1.
8 X APPROXIMATE TANDEM REPEATS, THR-RICH.
                     8 AA APPROXIMATE TANDEM REPEATS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels 247;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 662;
                                                                                                                                                              P-TYPE 2.
P-TYPE 3.
12 X APPROXIMATE TANDEM REPEATS,
        Glycoprotein; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   F085277F1ED2FD40 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tch 6.9%; Score 509.5; DB 1; al Similarity 28.1%; Pred. No. 1.2e-13; 223; Conservative 48; Mismatches 276;
                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                                                                                                                                                                                                                            SIMILARITY
                                                                                                                                                                                                                                                                                                                                                         SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                   BY SIMILARITY BY SIMILARITY
                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                       SIMILARITY
                                                                                                                                                                                                                                                                                                                             SIMILARITY
                                                                                                                                                                                                                                                                                                                                      SIMILARITY
                              ALA/THR-RICH
                                                                                                                                                                                                                                                                              P-TYPE 4.
P-TYPE 5.
P-TYPE 6.
                                                                                                                                                                                     THR-RICH.
                                                                                                                              2-4.
2-5.
2-6.
2-7.
2-8.
P-TYPE 2
                                                                                                                                                                                                                                                                                                                                                                                                                                ^ ^
                                                                                                                                                                                                                                                                                                                                                                                                                                              ۸
                                                         L-4.
L-5.
L-6.
                                                                                                                                                                                                                                                                                                                               BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                    MM;
  PROSITE; PS00025; P_TREFOIL; 6
Repeat; Amphibian skin; Glycopi
                                                                                                                                                                                                                                                                                                                                                                                                                                                    67774
                                                                                                                                                                                                                                                                                                                                                                       551
563
599
598
610
                                                                                                                                                                                            4411
4419
4431
4443
4472
4472
4472
5525
5525
5526
6611
1188
1188
1189
333
                                                                                                                                                   294
301
347
394
522
                                                                                                                                                                                                                                                                                                                                                                                                                                              415
662 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                            4402
4412
4420
4432
4444
4454
4403
4403
4403
620
620
1162
307
                                     DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                 DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                        DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                               DOMAIN
DOMAIN
DOMAIN
DISULFID
                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                      REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                           REPEAT
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                     REPEAT
REPEAT
                                                                                              DOMAIN
DOMAIN
                                                                                                                  REPEAT
REPEAT
                                                                                                                                      REPEAT
                                       REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                 REPEAT
                         DOMAIN
                                                                                                             REPEAT
                                                                                                                                 REPEAT
                                                                                                                                               REPEAT
                                                                                                                                                     REPEAT
                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
```

```
SEGUENCE FROM N.A.
MEDLINE-92178992; PubMed-154281;
MINITZERITH M., Acker J., Vicalire S., Vigneron M., Kedinger C.;
Wintrzerith M., Acker J., Vicalire S., Vignerase II largest subunit.";
"Complete sequence of the human RNA polymerase II largest subunit.";
"Lomplete Sequence of the human RNA polymerase II largest subunit.";
"Lomplete Acids Res. 20:910-910(1992).
                                                                                                                        01-MAR-1992 (Rel. 21, Created)
01-MAR-1992 (Rel. 21, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT (EC 2.7.7.6) (RPB1).
                                                                                                                                                                                                                                                                Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                        PRT; 1970 AA
                                                                                                            STANDARD;
567 LPOVADCKVAPSSR 580
```

267 ITNKQTSTDGKEKTISAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP 326

δy g

51

999

463

289

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Burpapan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41;
                                                                                                                                                                                                                                                                                                             Intermediate filament; Coiled coil; Neurone; Phosphorylation; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119 KKTKKVIESE--EITEEHSVS-ENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLK 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              176 VK------DNKKNRTKKRPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTT 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  205 VKSEEKIKVVE------KSEKETVIVEEQTEEIQVTEEVTEEDKEAGGEEEEE 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  282 SAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP----TPTTPKEPA-S 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               253 AEEGGEEAATTS-----PPAEEAASP----EKETKSPVKEEAKSPAEAKSPAEAKSPAEA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             304 KSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKS-PAEAKSPAEVKSPATVKSPGEAKSPAE 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              385 PKEPA----PTITKEPA----PTITKSAPTIPKEPA----PTIPKKPA----PTIPKEPA 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         429 ----PTTPKEP----TPTTPKEPAPTTKEP----APTTPKEPA----PTAPKKPAPTTPK 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      473 EPAPTTPKEPAPTTTKEPSPTTPKEPA----PTTTKS-----APTTTKEPAPTTTKS--- 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               337 TTPKEPTPTIKSAPTIPKEPA----PTTTKSAPTIPKEPA----PTTTKEP----APTT
                                                                                                                                                                                                                                                                                                                                                    REPEATS OF K-S-P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 831;
                                                                                                                                                                                                                                                                                                                                            51 X 3 AA TANDEM REPEATS OF R L -> I (IN REF. 2).
L -> S (IN REF. 2).
L -> T (IN REF. 2).
M -> T (IN REF. 2).
K -> N (IN REF. 2).
K -> N (IN REF. 1).
C -> E (IN REF. 1).
P -> S (IN REF. 2 AND 4).
A -> V (IN REF. 4).
A -> V (IN REF. 4).
A -> CST (IN REF. 4).
A -> L (IN REF. 2 AND 4).
A -> L (IN REF. 2 AND 4).
A -> CST (IN REF. 2 AND 4).
A -> V (IN REF. 2 AND 4).
A -> V (IN REF. 2 AND 4).
A -> P (IN REF. 2 AND 4).
B -> P (IN REF. 2 AND 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.0%; Score 517.5; DB 1; 26.6%; Pred. No. 7.2e-14; tive 76; Mismatches 345;
                                                                                                                                       EMBL; M37227; AAA41693.1; ALT_ERAME.
EMBL; X13804; CAA32038.1; ALT_ERAME.
EMBL; M21964; AAA41695.1;
EMBL; J04517; AAA41692.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MM;
                                                                                                                                                                                                                                                                  InterPro; IPR001664; IF.
Pfam; PF00038; filament; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 217; Conservative
                                                                                                                                                                                                                                                                                               PS00226; IF; 1.
                                                                                                                                                                                                                                                                                                                                               641
164
185
193
199
373
373
482
571
                                                                                                                                                                                                   PIR; A30796; A30796.
PIR; A25649; A25649.
PIR; B25649; B25649.
PIR; S02003; S02003.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       591
727
757
769
775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                 PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222
윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                           634
                                                                                                                                                                                              694
                                                                                                                                                                                                                                               754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C.1 (thm.v.l.) 110m Acurery
Dolymorphism.";
J. Biol. Chem. 267:24620-24624(1992).
-! FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL
INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.
-! SUBCELLULAR LOCATION: SECRETED.
-! ALTERNATIVE PRODUCTS: A NUMBER OF DIFFERENT FORMS OF THE PROTEIN MAY BE PROFOUCED BY ALTERNATIVE SPLICING.
-! TISSUE SPECIFICITY: SIRIN.
-! PTM: EXTENSIVELY O-GLYCOSYLATED.
-! SIMILARITY: CONTAINS 6 P-TYPE (TREFOIL) DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hauser F., Hoffmann W.; "P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
KPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTP
                                            --APTIPKEPSPTTTKEPA----PTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTK
                                                                         539 AKSPEKAKSPVKEEIKPPAEVKSPEKAKSPMRKEAKSPEKAKTLDVKSPEAKPPAKEEAK
                                                                                                                           TTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTTPKEPAPTTPKETAPTTPKGTAPTTLK
                                                                                                                                                                                                                                     EPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                    KGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoldea; Pipidae;
Xenopodinae; Xenopus,
                                                                                                                                                                                                                                                                                                                                                                  815 PKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDE 851
                                                                                                                                                                                                                                                                                                                                                                                               790 -- KGLPQEPSKPKTEKAEKSSTDQKDSQPSEKAPED 824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-1994 (Rel. 30, Created)
01-0CT-1994 (Rel. 30, Last sequence update)
101-0CT-1994 (Rel. 30, Last sequence update)
101-0CT-1994 (Rel. 30, Last annotation update)
Xenopus laevis (Afriçan Clewed frog)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-93077556; PubMed-1447205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000519; P_trefoil.
Pfam; PF00088; trefoil; 6.
SMART; SM00018; P; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; L02115; AAA74725.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , A45155.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR; A45155; A45155
HSSP; P04002; 1WFA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-8355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE-Skin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MUC1_XENLA
Q05049;
                                                                                                         575
                                                                                                                                                                        635
                                                                                                                                                                                                                                                                                                    755
                                                                                                                                                                                                                                      695
                                                                                                                                                                                                                                                                     707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MUC1_XENLA
             qq
                                          ò
                                                                       g
                                                                                                       ò
                                                                                                                                                                    δ
                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                              qq
```

```
54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    006
                                                             ADHESTVE PLAQUE MATRIX PROTEIN.
NONREPETIVE LINKER.
TANDEM REPEATS OF Y-K-[PS]-K-[IP]-[ST]-Y-P-[PS]-[ST]-[ST]-NONAPEPTIDE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Y---PPSYKPKKTYPPTXK-----PKISYPPTXK-----TKPSYPASYKRKTSYPPTXKPK 735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RITPETITAAPKMIKETAITIEKTIESKITATITQVISTITQDTIPFKITILKTITLAPK 960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                736 ISYPSTYKAKP----SYPPTYKPKPSYASSYKPKIRYPPTYKPKPSYASSYK----PK 785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EPSPITIKEPA-PITPKEPAPITPKKP-API-TPK------EPAPITPKEPAPITIKKP 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             411 T-YKPKISYPPTYKPKASYVSSYKSKKTYPPTYKPKLSYPPTYKPKPSYPPTYKPKITYP 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     739
                                                                                                                                                                                                                                                                                                                                                                                                         SAPT-TPKEPAPT--TIKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTT--PKK 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             K-PKPSYPPTYKPKITYPPT-YKPKPSYPTPYKQKPSYPPIYKSKSSYPTSYKSKKTYPP 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            631
                                                                                                                                                                                                                                                                                                                         85 YPTKRYQPTYGSKTNYPPIYKPIAKKLSSYKAIKTTYPAYKAKTSYPPSYK-----HKIT 139
                                                                                                                                                                                                                                                                                                                                                       363
                                                                                                                                                                                                                                                                                                                                                                                 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                471
                                                                                                                                                                                                                                           194 VVDEAGSGLDNGDFKVTTPDTST--TQH-----NKVSTSPKITTAKPINPRPSLPPNS 244
                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKT----TERDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        470 PTYKPKPSYPPTYKPKITYPPTYKRKPSYPTPYKOKPSYPPIYKSKSYPTSYKSKKTYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               530 PTYKPKITYPPTYK-----PKPSYRPSYKPKTTYPPTYKPKIRYPPTYKPKASYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          740 EPAPTIPKEPAPTIPKGIAPTILKEPAPTIP----KKPAPKELAPTITKGPISTISDKPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APTAPKEPAPTTPKETAP-----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --KEPAPTT--PKETAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      683 TT--PKGTAPTILKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPT-TPKGTAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTT - - PKETAPTTPKEPAPTTPK - - KPAPTTPET - - - - - - PPPTTSEVSTPTTKEPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29 VYGSAYSGASAGAYK-TLPGSHPYGSKHVPVYKPMNKIPT-PYI--SKKSYPAPYKPKGY
                                                                                                                                                                                                                                                                                                245 DTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVL
                                                                                                                                                                                                                                                                                                                                                     305 AKPTPKAETTTKGPALTTPKEPT-PTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                               140 YPPTYK-----PKITYP--PTYKOKPSYPPSYKPKTTYPPTYK-----PKITYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPT----TTKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYKPKITYPPTYKPKPSYPPSYKPKKTYSPTYKPKITYPPTYKPKPSYPPSYKPKTTYPP
                                                                                                                                                                                                                                                                                                                                                                                                                            PAPTTPKEPAPTT----PKEPTPTT--PKEPAPTTKEPAPTTPKEPAPT-APKKPAPTTP
                                                                                                                                                                                                                 Indels 181;
                                                                                                                                                                                       Length 872;
                                                                                                                                               98CC70D7C75FF3C4 CRC64;
                                                                                                                                                                               Score 518.5; DB 1;
Pred. No. 6.9e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PT-PTTPEEPAPTTPKAAAPNTPK-EPAPTTPKEPAPTTP---
                                                                                                                                                                                                                    80; Mismatches
                                                                                                                                      NONAPEPTIDE
                                                        POTENTIAL.
InterPro; IPR002965; P_rich_extensn. PRINTS; PR01216; ADHESIVEI. PRINTS; PR01217; PRICHEXTENSN.
                                                                                                                                                     Σ.
                                           Hydroxylation
                                                                                                                                                                                          7.0%;
                                                                                                                                                   101677
                                                                                                                                                                                                                       Conservative
                                                                                              872
                                                                                                                           192
                                                                    872
                                                                                                                                                     872 AA;
                                                                                                                                                                                                         Similarity
                                            Signal; Repeat;
                                                                    21
21
124
                                                                                                                           184
                                                                                                                                                                                                                        276;
                                                                                                                                                      SEQUENCE
                                                                                                                                                                                              Query Match
                                                                                                                                                                                                             Local
                                                                                    DOMAIN
                                                                                                   DOMAIN
                                                                                                                            DOMAIN
                                                                                                                                           DOMAIN
                                                          SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 901
                                                                                                                                                                                                           Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      964
                                                                                                                                                                                                                                                                                                                                                                                                                   364
                                                                                                                                                                                                                                                                                                                                                                                                                                             186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            584
                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΟŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                      ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΫ́
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
       DR
DR
DR
DR
DR
ET
FT
FT
FT
SO
                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                         δ
```

```
OF ANOMAL CALLERS.

OF ANOMAL CALLERS.

PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING OF THE LARGER NEURING PLAY. PLACE NEURING SON THE LEVELS (NF-M AND NF-H), THE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION. STRITES BELOWS TO THE INTERMEDIATE FILAMENT FAMILY.

CAUTION: REF. 1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 783 ONWARD AND IS LONGER DUE TO A FRAMESHIFT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTOC. Natl: Acad. Sci. U.S.A. 86:2463-2467(1989).

- PUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT.

- PTW. THERE ARE A NUMBER OF REPRATS OF THE RIPEPTIDE K-S-P, NFH IS PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE.
VTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ----KPTKAPKKP-TSTKKPK-T 1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŗ
                                      ---RYPPTYKPKP----SYASSYKPKIRYPPTYKPKPSYASSYKPKIT 827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Breen K.C., Robinson P.A., Wion D., Anderton B.H.; Partial sequence of the rat heavy neurofilament polypeptide (NF-H). Identification of putative phosphorylation sites."; FEBS Lett. 241:213-218(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       neurofilament peptide (NF-\dot{H}): developmental and tissue expression the rat, and mapping of its human homologue to chromosomes 1 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C., Felix J.M., Nussbaum J.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           large neurofilament subunit (NF-H) of the rat: cDNA cloning
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Robinson P.A., Wion D., Anderton B.H.;
"Isolation of a cDNA for the rat heavy neurofilament polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ,
D
                                                                                                                                                                                                                                                                                                                                                                                                              Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 318-831 FROM N.A.
MEDLINE=80184647; PubMed=2928342;
Lieberburg I., Spinner N., Snyder S., Anderson J., Goldgaber D., Smulowitz M., Carroll Z., Fananuel B.S., Breitner J., Rubin L.;
Smulowitz M., Carroll Z., Fananuel B.S., Breitner J., Rubin L.;
"Cloning of a cDNA encoding the rat high molecular weight
                                                                                                                                                                                                                                                       P16884; Q63368;
01-AUG-1990 (Rel. 15, Created)
01-FBE-1994 (Rel. 28, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA,NEUROFILAMENT PROTEIN)
(NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Biochem. Biophys. Res. Commun. 154:1099-1106(1988)
                                                                                                                        828 YPPTYKPKISYPPTYKPKITYPPTYKPKISYPPAYKPKISYPSQ
                                                                                   1016 MPRVRKPKTT--PTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQ
                                                                                                                                                                                                                                    ¥
                                                                                                                                                                                                                                    831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-89 AND 243-313 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dautigny A., Pham-Dinh D., Roussel Jolles P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=89065087; PubMed=3143606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PubMed-2457365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE=87080760; PubMed=2878828;
                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (NF-H).";
FBBS Lett. 209:203-205(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 37-831 FROM N.A.
                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                     Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             in situ detection.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=88309090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                    NEFH OR NFH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The
            961
                                                     786
                                                                                                                                                                                                                                            g
                                                                                                ò
                                                                                                                                     g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Bundermatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41;
 .. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
ENCYTON MIGHT FUNCTION AS A CALCIUM-SEQUESTERING "SPONGE" TO REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT BINDS 0.3 MOLOF PORTEIN.
OF CA-2 PER MOL OF PROPIEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APTSKVLAKP-TPK---AETTTKGPALTTPKEPTFTTPKEPASTTPKEPTFTTKSAPTT 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354 PKEPAPTITKSAPTTP-KEPAPTTTKEPAPTTPKEPAPTTTKEPAPT-----TTKSAPTTP 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 VIIPAPAPIAAASVTPVASVAPPVVAAPTP----PAASPVSTPVAVAQIPVAVSAPVAP 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117 PVAATPTPVVQIPVAAPVIATPPVAASAPTPAAVTPVISPVIASPPVVPANTTVPVAAPV 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             297 ETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAETPEVASVAVAETTPPVVPPV 356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPVSAPVAAPVTPSAVAAPVQVVSPAAVAP-----APAAPIAVTPVAPPPTLASVQPAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  660 TPKEPAPTTPKEPAPTTPKETAPTTP-KGTAPTTLKE------PAPTTPKKPAPKE 708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 409 KEPAPTTPKKPAP-----TTPKEPAPTTPKEPTP-----TTPKEPAPTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         455 PKEPA--PTAPKKPAPTTPKEPAPTTPKEPAPTTTKE-PSPTTPKEPAP-----TTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTTTKEP---APTTTKSAPTTPKEPSPTTTKEPA-PTTPKEPAPTTPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                563 TTP-----KEPAPTTTKKPAPTAPKEP---APTTPKET-----APTTPKKLTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            604 TPEKL-APTTPEKPAPTTPEELAPTTPEEPTPTTPEEP---APTTPKAAAPNTPKEPAPT
                                                                      -1- SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPARTMENT.
-1- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS OF COMPOUND EYES AND OCELLI.
-1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.2%; Score 530; DB 1; Length 865; 26.2%; Pred. No. 2.5e-14; .ive 73; Mismatches 433; Indels 152;
                                                                                                                                                                                                                                                                                                                                                                   A -> AVADAVVA (IN REF. 2).
I -> T (IN REF. 2).
T -> A (IN REF. 2).
T -> A (IN REF. 2).
P -> PP (IN REF. 2).
VQ -> AP (IN REF. 2).
I -> V (IN REF. 2).
S -> T (IN REF. 2).
A -> E (IN REF. 2).
I -> V (IN REF. 2).
C -> T (IN REF. 2).
A -> E (IN REF. 2).
V -> L (IN REF. 2).
V -> L (IN REF. 2).
D -> E (IN REF. 2).
W -> L (IN REF. 2).
D -> E (IN REF. 2).
W -> L (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ::
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84781 MW;
                                                                                                                                                                                                                                                                                      EMBL; L02111; AAA28405.1; -.
EMBL; L05080; AAA28420.1; -.
PIR; A47282; A47282.
                                                                                                                                                                                                                                                                                                                                   FBgn0010218; Cpn.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                   36
43
64
64
100
1127
1154
1160
534
699
703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       699
703
721
865 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                        DEVELOPMENT
                                                                                                                                                                                                                                                                                                                                                     Calcium-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             233;
                                                                                                                                                                                                                                                                                                                                                                CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                   FlyBase;
                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local
Matches 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
 Proc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      507
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                         968 ITTTEIMNKPEETAKPKDRATNSKATTPKPOKPTKAPKKPTSTKKPKTMPRVRKPKTTPT 1027
                                        109 LAPITIKEPISITSDKPAPITPKGTAPTTP-----KEPAPITPKEP--APITPKGTAPI 760
414 V----LPPVAAEPVPAVVAEETPETPAPASAPVTIAALDIPEVAPVIAAPSDAPAEAPSA 469
                                                                                                                                                        528 APPEAAADLIIEPVEPPAPIPDLLEQTISVPAVEAAESISS--PIPETSLPPPNEAVASP 585
                                                                                                                                                                                                                                                                                                                                                                                          698 DIAIPVIDPPVPQEIAVAEIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAA 757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11. WOL. EVOL. 43:348-356(1996).

1- MOL. EVOL. 43:348-356(1996).

1- FUNCTION: PROVIDES ADHESYVENESS TO THE MUSSEL'S FOOT. MUSSELS ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS ADHESIVE DES ADHESIVE THE MUSSEL'S ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIRROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

1- SUBCELLULIAR LOCATION: SECRETED.

1- TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

1- DOMAIN: ALMOST EXCLUSIVELY COMPOSED OF REPEATS OF A DECAPEPTIDE.

1- PTM: PROLINES IN THE REPEAT ARE HYDROXYLATED (SINGLE OR DI.) AND ALSO TYROSINE (THUS PRODUCING DOPA = 3,4-DIHYDROXYPHENYLALANINE).
                                                                                                                                                                                                                                                   641
                                                                                                                                                                                                                                                                                       914
                                                                                                                                                                                                                                                                                                                               697
                                                                                                                                                                                                                                                                                                                                                          915 KETATTTEKTTESKITATTTQVTST-----TTQDTTPFKITTLKTTTLA-PKVTTTKKT 967
                                                                                                                    761 TLKEPA-----PTTPKKPAPKELAPTTT----KGPTSTTSDKPAPTT----PKETAPTT
                                                                                                                                                                                                PKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPD-------ESTPEL
                                                                            AAPIVSTPP--TTASVPETTAPPAAVPTEPIDVSVLSEAAIETPVAPPVEVTTEVAVADV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Adhesive protein cDNA sequence of the mussel Mytilus coruscus and
                                                                                                                                                                                                                              857 SAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTP--ETTTAAPKMT
                                                                                                                                                                                                                                                                                                        30-MAY-2000 (Rel. 39, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ADHESIVE PLAQUE MATRIX PROTEIN PRECURSOR (FOOT PROTEIN 1) (MCFP1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mytilus coruscus (Sea mussel).
Sukaryota, Metazoa, Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
Mytiloidas; Mytilidae; Mytilus.
NCBI_TaxID=42192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Inoue K., Takeuchi Y., Takeyama S., Yamaha E., Yamazaki F., Odo
Harayama S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1028 PRKMTSTMPELNPTSRIAEAM---LQTTTRPNQTPNSKLVEVN-----PKS 1070
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             814 LRDLQTTDVSLLAIAATLDAIGEKLKDQKARNQOVMDRLCEIEKILGPPKS 864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       872 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-96394686; PubMed-8798340;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last seq.
20-AUG-2001 (Rel. 40, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D63777; BAA09850.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE=Foot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FP1_MYTCO
                                                                                                                                                                                                807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FP1_MYTCO
                                        ò
                                                                            g
                                                                                                                  ò
                                                                                                                                                        pp
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

InterPro; IPR002964; Adhesive_plag.

327 TPTTPKE-PASTTPKEPTPTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTP 385

```
QQ
                                                                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                        Q_{\mathbf{y}}
                                                                                                                                              qq
                                                                                                                                                                              Q_{\lambda}
                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                    δy
                   q
                                                        Óγ
                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEVELOPMENTAL STAGE: MAXIMAL ACTIVITY IN TRYPOMASTIGOTES, MINIMUM IN EPINASTICOTES AND NO DETECTABLE IN AMASTIGOTES.
MISCELLANEOUS: THE VARIABLE LENGTHS OF THE LONG TANDEM REPEAT
                                                                                                                                                                                                    Prioli R.P.;
"The Trypanosoma cruzi neuraminidase contains sequences similar to bacterial neuraminidases, YWTD repeats of the low density lipoprotein receptor, and type III modules of fibronectin.";
J. Exp. Med. 174:179-191(1991).
                                                                           Trypanosoma cruzi.
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
NCBI_TaxID=5693;
01-NOV-1991 (Rel. 20, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
SIALIDASE (EC 3.2.1.18) (NEURAMINIDASE) (NA) (MAJOR SURFACE ANTIGEN)
                                                                                                                                                                                                                                                                                                                                                                                            PARSITE INVASION OF CELLS.

CATALYTIC ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC LINRAGES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEURAMINYL RESIDUES TO GALACTOSE, N- ACETYLHEXOSAMINE, OR N- OR O-ACYLATED NEURAMINYL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS, GLYCOLIPIDS OR COLOMINIC ACID.

SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BNR 1.
BNR 2.
BNR 2.
FIBRONECTIN TYPE-III.
44 x 12 AA TANDEM REPEATS, LTR DOMAIN.
                                                                                                                                                                                                                                                                                                                                   of
                                                                                                                                                                                                                                                                                                                                                           trypomastigotes.";
trop. wed. parasitol. 42:146-150(1991).
-1- FUNCTION: DEVELOPMENTALLY REGULATED NEURAMINIDASE IMPLICATED IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN COULD ACCOUNT IN PART FOR THE POLYMORPHISM OF THE TNCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. ) (POTENTIAL).
(GLCNAC. .) (POTENTIAL).
(GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                    MEDLINE-91376547; PubMed-1896773;
Prioli R.P., Mejia J.S., Aji T., Aikawa M., Pereira M.E.A.;
"Trypanosoma cruzi: localization of neuraminidase on the surface
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES. SIMILARITY: CONTAINS 3 BNR REPEATS.
                                                                                                                                         SEQUENCE FROM N.A.
STRAIN-SILVIO X-10/4;
MEDLINE-912709; PubMed-1711561;
Pereira M.E.A., Mejia J.S., Ortega-Barria E., Matzilevich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase; Glycosidase; Glycoprotein; Repeat; GPI-anchor;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07049221897C6A40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CYS-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             120032 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL, M61732; AAA30255.1; -. PIR, JH0557; JH0557. HSSP; P29768; IDIL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002860; BNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          457
34
174
220
220
588
1120
342
394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF02012; BNR; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                            SUBCELLULAR LOCATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Phosphorylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POSSIBLE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                    TCNA.
```

```
STRAIN-CANTON-S;
MEDLINE-93165729; Pubmed-8094559;
MATLIN J.H., Benzer S., Rudnicka M., Miller C.A.;
"Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";
Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila melanogaster (Fruit fly).

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pteryyota: Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea: Drosophilidae; Drosophila.

NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                   712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          772 KPAPKELAPITIKGPISITSDKPAPITPK----EIAPTIPKEPAPITPKKPAPITPEIPP 827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      886
                                                                                                                                                                                                                                                                                                                                                                                                              713 TTKEPTSTTSDKPAPTTPKGTAPTTPKE-PAPTTPKEPAPTTPKGTAPTTLKEPAPTTPK 771
                                                                                                                                                                                                                                                                                       652
                                                         652 ADSSAHGTPSTPVDSSAHSTPSTPVDSSAHGTPSTP 711
                                                                                                                        491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=9315730; PubMed=8434015;
MEDLINE=93155730; Xue N., Harshman K.D.;
Ballinger D.G., Xue N., Harshman K.D.;
"A Drosophila photoreceptor cell-specific protein, calphotin, binds
calcium and contains a leucine zipper.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      828 PTTSEVSTPTTTKEPTTIHKSPDESTP-ELSAEPTPKALENSPKEPGVPTTKTPAATKPE
                                                                                                                                                                                                                                                                                                                                                   PKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT
                                                                                                                                                                                 548 TP-KKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKL----TP
                                                                                                                                                                                                                                               603 TTP-EKLAPTTPEKPAPTTPEELAPTTPEE----PTPTTPEEPA----PTTP-KAAAPNT
                                                                                                                                                                                                                                                                                                          936 PSTPADSS----AHSTPSTPADSSAHST-PSTPADSS----AHSTPSTPVDSSAHST
                                                                                                   KEPTP-TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTKEPS
                                                                                                                                                               PITPKEPAPITIKSAPITIKEPAPITIKSAPITPKEPSPITIKEPAPITPKEPA----PT
KE-PAPTITKEPAPITIKSAPITPKE-PAPTIPKKP-----APTIPKEPA---PTIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-0CT-1993 (Rel. 27, Created)
01-0CT-1993 (Rel. 27, Last sequence update)
01-FEB-1994 (Rel. 28, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-CANTON-S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1126 GT 1127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           887 MT 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CALPHOTIN.
CPN OR CAP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CPN_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10
                                                                                                                                                                                                                                                                                                                                                           653
                                                                                                                                                                                                                                                                                                                                                                                                                                                     984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CPN_DROME
                                                386
                                                                                                             433
                                                                                                                                                                        492
```

34;

Gaps

Indels 113; Length 1162;

DB 1; 212;

7.2%; Score 530.5; DB 29.4%; Pred. No. 3e-14;

Ouery Match 7.2%; Score 530.5; I Best Local Similarity 29.4%; Pred. No. 3e-14Matches 177; Conservative 100; Mismatches

```
F. S
                                                                                                                   ò
                                                                                                                                     ద
                                                                                                                                                        à
                                                                                                                                                                       qq
                                                                                                                                                                                            ò
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                              òγ
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                  OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 534
TO 716 AND IS SHORTER DUE TO FRAMESHIFTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLU-RICH (ACIDIC).
50 X 6 AA TANDEM REPEATS OF K-S-P-A-E-A.
GLU/LYS-RICH.
               SEQUENCE FROM N.A.
MEDLINE-89089138; PubMed-3145094;
Shnetdman P.S., Carden M.J., Lees J.F., Lazzarini R.A.;
"The structure of the largest murine neurofilament protein (NF-H)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Neurone; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   K -> OA (IN REF. 2 AND 3).
A -> AR (IN REF. 2 AND 3).
S -> T (IN REF. 2 AND 3).
L -> G (IN REF. 2 AND 3).
P -> PREAKSP (IN REF. 3).
MISSING (IN REF. 3).
G -> A (IN REF. 3).
V -> M (IN REF. 2 AND 3).
                                                       revealed by cDNA and genomic sequences.";
Brain Res. 464:217-231(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LINKER 12.
COIL 2A.
LINKER 2.
COIL 2B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LINKER 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COIL 1B.
                                                                                                                                                                                                                                                                                                                                                          EMBL; M24496; AAA39813.1; -...
EMBL; M23449; AAA39813.1; JOINED.
EMBL; M24494; AAA39813.1; JOINED.
EMBL; M24495; AAA39813.1; JOINED.
EMBL; M35131; AAA39809.1; ALT_FRAME.
EMBL; Z31012; CAA83229.1; -...
                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001664; IF.
Pfam; PF00038; filament; 1.
PROSITE; PS00226; IF; 1.
Intermediate filament; Colled coil;
                                                                                 SEQUENCE FROM N.A.
STRAIN-SWISS WEBSTER; TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              141
233
261
261
283
287
408
133
199
281
Gene 68:307-314(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; JT0368; OFMSH.
PIR; A43778; A43778.
MGD; MGI:97309; Nfh.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Repeat.
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
```

```
43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 KKTKKVIESE--EITEEHSVS-ENQESSSSSSSSSSTIWKIKSSKNSAANRELQKKLK 175
                                                                                                                                                                  433 VKSEEMIKVVE------KSEKETVIVEGQTEEIRVTEGVTEEEDKEAQGQEGEE 480
                                                                                                             Gaps
                                                                                                                                                                                                                VK-------DNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        592 SPSEAKS-PAEAKSPAEAKSPAEAKSPAEAKSPAEAKSPAEAKSPAEAKSPATVKSPGEAKSPSEA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           604 TPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEE-PAPTTPKAAAPNTPKEPAPTT-- 660
                                                                                                                                                                                                                                                                                        222 VSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTT 281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                441 КЕРАРТТКЕРАРТТРКЕРА----РТАРККР----АРТТРКЕРАРТТРКЕРАРТТКЕРSР 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            935
                                                                                                                                                                                                                                                                                                                                                                                                                                                            282 SAKETQSIEKTSAKOLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKE
                                                                                                                                                                                                                                                                                                                                                                                      493 TIPKEPA----PITTKSAPTTTKEPAPTTTKSAPTTPKE-PSPTTTKEPA----PTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 342 РТРТІКЅАРТТРКЕРАРТТКЅАРТТРКЕР-АРТТКЕРАРТТРКЕРАРТТКЕР----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          397 APTITKSAPTIPKEPA----PTIPKKPA----PTIPKEPA----PTIPKEP----TPTIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  820 PPEKAKPLDVKSPEAQTPVQEEATVPTDIRP--PEQVKSPAKEKAKSPE--KEEAKTSEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --PKEPAPTTP-----KEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITKEPISITSD--KPAPITPKGTAPTTPKEPAPITPKEPAPITPKGTAPITLKEPAPIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 770 РККРАРКЕLАРТТІКGРІSІТІЗОКРАРІТРКЕТАРІТРКЕРАРІТРККРАРІТРЕТРЕР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    544 PAPTTPKKPAPTTPKEPAPTTPKEPAPTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        830 TSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTT
                                                                                                           Indels 186;
                                                                         Length 1087;
   -> N (IN REF. 2 AND 3), 57BAC76A38EDICB9 CRC64;
                                                                                          Pred. No. 4.6e-15;
96; Mismatches 358;
                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1162 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            890 TAKDKTTERDLRTTPETTTAAPKMTKETATTTEK 923
                                                                       7.4%; Score 551;
26.8%; Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
843 T. -
116612 MW;
                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
843 84
1087 AA;
                                                                                        Similarity
                                                                                        Best Local Sim
Matches 234;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TCNA_TRYCR
P23253;
CONFLICT
                                                                       Query Match
                                                                                                                                                                                                                  176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TCNA_TRYCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A D
```

641DD2278AB28524 CRC64;

65406 MW;

620 AA;

SEQUENCE

S

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PENTAPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            H-A-P-P.
H-A-P-P.
2 X 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.
                             8.14
                                                           691
                                                                                       855
                                                                                                                     745
                                                                                                                                                                              PKEPAPITPKGTAPITLKEPAPITPKKPAPKELAPITIKGPISTISDKPAPITPKET--A 803
                                                                                                                                                                                               631
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Viridianiania, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnollophyta, eudicotyledons, core eudicots, Asteridae, euasterids I; Solanales, Solanaceae, Nicotiana.
                 KEPAPTTPKETAPT-----TPKKLTPTTPEKLA-----PTTPEKPAPTTPEELAPTTPEE
                                                                                                                     TLKEPAPTTPKKPAPKELAPTTTKEPTSTT-----SDKPAPTTPKGTAPTTPKEPAPTT
                                                                                                                                      PTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPT
                                                                                       815 KPTISPEKPTISTEK------PTIPTE-KPTIPTE------ETTISTEKLTIPT
                                                                                                                                                                                                                                                                                                                                                                           01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-MAR-1992 (Rel. 21, Last annotation update)
EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN=CV. XANTHI: TISSUE=Leaf;
MEDLINE=90128263; PubMed=2612909;
MEDLINE=90128263; PubMed=2612909;
Rellor B., Lamb C.J.;
"Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation.";
Genes Dev. 31639-164(1989).
-i- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2. CONTAINS THE SER-PRO(4) REPEATS.
3 X APPROXIMATE TANDEM REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PER
SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
GLYCOSYLATED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Structural protein;
                                                                                                                                                                                                                                                                                                                                               620 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; S06733; S06733.
Repeat; Cell wall; Glycoprotein; Signal;
Hydroxylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EXTENSIN.
                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nicotiana tabacum (Common tobacco)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X13885; CAA32090.1; -.
                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              620
73
242
242
620
620
                                                                                                                                                                                                                                                              | : | | |:
967 PASCKSPRPS 976
                                                                                                                                                                                                                                           804 PTTPKEPAPT 813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4097;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              70
148
2229
229
229
236
499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MAIN ROOT
                                                                                                                                                                                                                                                                                                                                               EXTN_TOBAC
P13983;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                  EXTN_TOBAC
                                758
                                                                                                                                                     856
                                                                                                                                                                                 746
                                                               632
                                                                                                                                                                                                                                                                                                                                                               δ
                                d
                                                              δ
                                                                                        g
                                                                                                                         δy
                                                                                                                                                  g
                                                                                                                                                                                   δ
                                                                                                                                                                                                             d
                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                       g
```

```
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      т.
.
                                                                                                                                                            SHGHLPPSVGGPPPHRGHLPPSRGFNPPPSPSHPPPSYGAPPPSHGPGHLPSHGQR 140
                                                                                                                                                                                                                                520
                                                                                                                                                                                                                                                                                              681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         370 PPPPSSSPPPSFSPPPPTYEQSPPPPPAXSPPLPARPTYSPPPT--YSPPPPTYAQPP 427
                                                                311 AETTIK-GPALTIP--KEPIPIT----PKEPASTIPKE---PTPIIKSAPIIPKEPAP 359
                                                                                                                                                                                                                                                                                                                                                                                                               624
                                  Gaps
                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     482 PPPSPIYSPPPQVQPLPPTFSPPPRRIHLPPPPHRQP---RPPTPTYGQPPSPPTFSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       742 APTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         802 TAPTTPKEPAPT - - TPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       539 РРРКОІНЅРРРРИМОРКТРТРТУGOPPSPPTFSAPPPROIHSPPPPROPREDPTPTYGOP
                                                                                                                                                                                                                                                                                                                                                                                                               KPAPT ---APKEPAPTTPKETAPTTPKKL-TPTT-----PEKLAPTTPEKPAPTTPEEL
                                                                                                                                                                                                                                                                                                                                                                                                                                         310 PPTPTPTFSPPPPAXSPPPPTXSPPPPTXLPLPSSPIXSPPPPVXSPPPPTXL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APTTPEEPTPTTPEEPAPTTPKAAAP----NTPKEPAPTTPKEPAPTTPKEPAPTTPKETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    428 P-----LPPTYSPPPAYSPPPPTYSPPPPTYSPPPAYAQPPPPTYSPPPAYSP
                                                                                                                                                                                                                                                                                                                                                                   250 RHLPPSPRRQPQPPTYSPPPAYAQSPQPSPTYSPPPPTYSPPPSPTYSPPPAYSPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    682 PITPKGTAPITLKEPAPITPKKPAPKELAPITTKEPISITSDKPAPITPKGTAPITPKEP
                                                                                                    24 AEATTOYGGYLPPPVTSOPPPSSIGLSPPSAPTTTPPSRGHVPSP---RHAPPRHAYPPP
                                                                                                                                        TITKSAPTIPKEP-----APTITKEPAPI----TPKEPAPITIKEPAPITIKSAPTIPKE
                                                                                                                                                                                                          PAPTTPKKPAPT - - - TPKEPAPTTPKEPTPTFKEPAPTTK - - EPAPTTPKEPAPTKK
                                                                                                                                                                                                                                                                           PAPITPKEPAPITIKEPAPITIKEPSPTTPK---EPAPITIKSAPTITKEPAPITIKS--
                                                                                                                                                                                                                                                                                                                                           ---APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPT----TPKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Sequence and structure of the mouse gene coding for the largest neurofilament subunit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P19246; Q61959;
01-NOV-1990 (Rel. 16, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT PROTEIN)
NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H).
                                  88;
Length 620;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINES-89121513; PubMed-3220257;
Julien J.-P., Cote F., Beaudet L., Sidky M., Flavell D.,
Mushynski W.;
                                  64; Mismatches 304; Indels
Score 551; DB 1;
Pred. No. 2.9e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 1087 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             860 PTPKALENSPKEPGVPTTKTPAATKP 885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             599 PSP-----PTTYSPPSPPP
 7.48; 27.28;
                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (Mouse)
                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                     Matches 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NFH_MOUSE
   Query Match
Best Local S
                                                                                                                                          360
                                                                                                                                                                                                              411
                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                              466
                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     625
                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                  521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NFH_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                            С
                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                            Оþ
                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                        g
                                                                                                                                           ò
```

```
RESULT EAN_HUMN ID ZNAN_HUMN ID ZNAN_HUMN ID ZNAN_HOW OF AN ON HOW OF EAN.

OS HOW OOK HOW OOK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09Y493; 0U0Z18;
20-AUG-2001 (Rel. 40, C
20-AUG-2001 (Rel. 40, I
20-AUG-2001 (Rel. 40, I
                                                EMBL; AF053356; AAC78790.1; -. EMBL; U83191; AAC51208.1; -. MIM; 602372; -.
                                                                                                                                                                     the European Bioinformatics Institute. The use by non-profit institutions as long a modified and this statement is not removed. entitles requires a license agreement (See )
                                                                                                                                                                                                                                                                                       between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           zonadhesin gene (ZAN).";
Genomics 41:119-122(1997).
-i- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZAN_HUMAN
Q9Y493; O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97271566; PubMed-9126492;
Gao Z., Harumi T., Garbers D.L.;
"Chromosome localization of the mouse zonadhesin gene and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    reveals 17 genes.";
Genome Res. 8:1060-1073(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Large-scale sequencing of two regions in human chromosome 7q22 analysis of 650 kb of genomic sequence around the EPO and CUTL1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 2338-2700 FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Tsu1 L.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Glockner G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-99018118; PubMed-9799793;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-2379 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-9606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZONADHESIN (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         849 YNKKISYPSSYKAKTSYPPAYKPTNR 874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        913
                                                                                                                                                                                                                                                                                                                                          DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS, THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF SPERMATOZOA OR PROMOTING ADHESTON TO THE CVIDICTAL ISTHMUS. DOMAIN: THE WIFD DOMAINS 2 AND 3 MAY MEDIATE COVALENT OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2). SIMILARITY: CONTAINS 3.5 MAM DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIGNALING.
SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN,
APICAL REGION OF THE SPERM HEAD (BY SIMILARITY)
TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPP
DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM,
ZONA PELLUCIDA.
                                                                                                                                                                                                                SWISS-PROT entry is copyright. It is produced through een the Swiss Institute of Bioinformatics and the Enteropean Bioinformatics Institute. There are no restriby non-profit institutions as long as its content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --YKPKLTYPPTYK-----PKPSYPPSYKPKITYPSTYKLKPSYPPTYKSKTSYPPT 848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IMNKPEETAKPKDRATNSKATTPKPQKPTKAPKK---PTSTKKPKTMPRVRKPKTTPTP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -TYKAKPTYPPTYKAK------PSYPPTYKPKPSYPPTYKSKSIYPSSYKPKKTYPPT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE
                                                                                                                                                  an email to license@isb-sib.ch)
IPR000561; EGF-like IPR000998; MAM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rosenthal A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scherer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---RKMTSTMPELNPTSR 1043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S., Schattevoy R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation updat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Z
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2700
                                                                                                                                                             (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Boright A., Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ⋛
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MANNER TO THE ZONA RECOGNITION AND/OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hominidae;
                                                                                                                                                                                             Usage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HAPLOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADHESION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EXCLUSIVELY ON
                                                                                                                                                                                                by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo
                                                                                                                                                                                                                                           restrictions
                                                                                                                                                                                                and for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPERMATIDS
                                                                                                                                                                                                                                                                         EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CUTL1 loci
                                                                                                                                                                                                                                                                                                  Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TO THE
                                                                                                                                                                                                                                                                      collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PELLUCIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  human
                                                                                                                                                                                           n no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1028
                                                                                                                                                                                                                      n 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         798
                                                                                                                                                                                                                                           its
```

```
δõ
                                                                      Db
     밁
                                     В
                                                   ρ
                                                                                      οy
                                                                                                        밁
                                                                                                                      ρ
                                                                                                                                          Вb
                                                                                                                                                         Š
                                                                                                                                                                          В
                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                      Query Match
Best Local
                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                    CONFLICT
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Glycoprotein;
NON_TER
DOMAIN <
     702
                    527
                                      648
                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PROSITE; PS01186; EGF_2; PROSITE; PS00740; MAM_1; PROSITE; PS50060; MAM_2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                     467
                                                                       597
                                                                                       407
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRINTS;
                                                                                                        540
                                                                                                                                          488
                                                                                                                                                          293
                                                                                                                                                                          445
                                                                                                                                                                                           233
                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Piam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                  Local Similarity
   EKPTISTEE-
                                                 APTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPK
               EPSPTTTKEPAPTTPKEPAPTTPKKPA----PTTPKEPAPTTPKEPAPTTTKKPAPTAP
                                                                                 TPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKP
                                                                                                   TEKPTIPSE-KPTIPSEKPTISTEKPTVPTEE---PTTPTEETTTYMEEPVIPTEKPSIPT
                                                                                                                                                                       PVKVLPELPPVSPVSS----TGPSETTGLTENPTISTK----
                                                                                                                                                                                     PINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKT 292
                                -PTTPTE-KPTIPTEKPTISTEKPTIPTEK-PTISPEKLTIPTEKLTIPT---EKPTIPT
                                                                 --EKPSIPTEK----PTISMEETIISTEKPTICPEKPTIPTEK----PTIPTEKSTISPEK-
                                                                                                                   KSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA-PTTTKSAPT
                                                                                                                                    SVTTEKPT-----VPKEKPTIPTEKPTISTEK----PTIPSEKPNMPSEKPTIPSEKPTIL
                                                                                                                                                   SAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPT-----PTTI 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; PF00629; MAM; 3.
; PF01826; TIL; 4.
; PF02345; TILa; 4
                                                                                                                                                                                                            199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF00094; vwd; 4.
S; PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SM00137; MAM; 2.; SM00214; VWC; 1.; SM00011; VWC_def; 3; SM00216; VWD; 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002965; P_rich_extensn.
IPR002919; TIL.
IPR003328; TILa.
IPR001007; VWRC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR001846;
                                                                                                                                                                                                                                                                     403
1023
1099
1618
1737
1832
1878
2136
2505
                                                                                                                                                                                                                                                                                                                                                                                  953
1066
1455
1862
                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                            1
112
161
322
483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane;
-PTTPTEETTISTEKPSIPMEKPTLPTEETTTSVEETTISTEKLTIPM--
                                                                                                                                                                                                                                                     AA;
                                                                                                                                                                                                                                                                                              403
1023
1099
1618
1737
1832
1878
                                                                                                                                                                                                                7.5%;
                                                                                                                                                                                                                                                                                                                                                                         2684
                                                                                                                                                                                                                                                                                                                                                                                                                            109
136
326
446
951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Vwd
                                                                                                                                                                                                                                                    293013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ω
                                                                                                                                                                                                          64;
                                                                                                                                                                                                                                                    MW;
                                                                                                                                                                                                                                                                           N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                 Score 555.5; DB 1
Pred. No. 6.4e-15;
                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                        VWFD
                                                                                                                                                                                                                                                                                                                                                                                66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
(MUCIN-LIKE DOMAIN).
VWFD 1 (PARTIAL).
VWFD 2.
VWFD 3.
VWFD 4.
                                                                                                                                                                                                                                                                   NNQKMA
                                                                                                                                                                                                                                                                                                                                                                                                                                    MAM
                                                                                                                                                                                                                                                                                                                                                                                                                                             MAM
MAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cell adhesion; Repeat
                                                                                                                                                                                                                                                                                                                                                                 EGF-LIKE
                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                  80E60CC0B12277B1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (PARTIAL).
                                                                                                                                                                                                                                                                   ť
                                                                                                                                                                                                                                                                           (GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
                                                                                                                                                                                                                                                                   RAGPGP
                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                         240;
                                                                                                                                                                                                                                                                   (IN
                                                                                                                                                                       ------KPTVSIEKP 487
                                                                                                                                                                                                                                                                 REF.
                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                   (POTENTIAL).
                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                         2700;
                                                                                                                                                                                                       107;
                                                                                                                                                                                                      Gaps
                                                                                  466
                                                                                                                    406
                                                                                                                                    539
757
                                701
                                                 526
                                                                   647
                                                                                                   596
                                                                                                                                                                                                       32;
```

```
RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           멍
REPERENCE OF CONCORD CONTRACTOR STANDARD CONTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                             30-MAY-2000
30-MAY-2000
20-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FP1_MYTED
Q25460;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIINE-91025829; PubMed=1367451; Filpula D.R., Lee S.M., Link R.P., St. "Structural and functional repetition
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FP1
                                                                                 modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=6550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mytiloidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADHESIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Waite J.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Structural and protein.";
                  EMBL; X54422; CAA38294.1; -
InterPro; IPR002964; Adhesi
                                                                                                                                                                              the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=83135732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PARTIAL SEQUENCE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biotechnol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                       s SWISS-PROT entry is copyright. It is produced through a collaboration ween the Swiss Institute of Bioinformatics and the EMBL outstation - Buropean Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HNVTSASGSASGSASTLVHNGTSARATTTPASKSTPFSIPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   \tt QT-PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPML
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     \tt PGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TQDTTP--FKITTLKTTTLAPKVTTTKKTITTT--EIMNKPEETAKPKDRATNSKATTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STAPPAHGVTSAPDTRPAPGSTAPPAH---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P--GSTAPPAHGVTSAPDTRP---APGSTAPPAHGVTSAPDTRPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --PAHGVTSAPDTRPAPGSTAPPAHGVTSAP----DTRPAPGSTAPPAHGVTSAPDTRPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTAKDKTTERDLRTTPETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SDETNICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPS 1156
                                                                                                                                                                                                                                                                  DOPA) DERIVED FROM TYROSINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  edulis (Blue mussel).
ta; Metazoa; Mollusca; Bivalvia;
dea; Mytilidae; Mytilus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             000 (Rel. 39, Created)
000 (Rel. 39, Last sequence update)
001 (Rel. 40, Last annotation update)
PLAQUE MATRIX PROTEIN (POLYPHENOLIC
                                                                                                                                                    non-profit institutions as long
    IPR002964;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MEFP1) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Prog.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AND POST-TRANSLATIONAL PubMed=6298211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6:171-177(1990)
    Adhesive_plaq.
P_rich_extensn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----TAAPKMTKETATTTEKTTESKITATTTQVTSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Strausberg S.L., Strausberg ion in a marine mussel adhes:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GVTSAPDNRPALGSTA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               875
                                                                                                                                                           There are no restrictions ng as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pteriomorphia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MODIFICATIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADHESIVE PROTEIN) (FOOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mytiloida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              adhesive
                                                                                                                             is
for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -PPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ġ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R.L.;
                                                                                                                                             in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    955
                                                                                                                                                                                       1 ts
```

```
DR FT FT SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                QΥ
                                                                                                                                                                                                                                                                                                                         망
                                                                                                                                                                                                                                                                                                                                             QY
                                                                                                                                                                                                                                                                                                                                                                     Вb
                                                                                                                                                                                                                                                                                                                                                                                          VΩ
                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DЬ
                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                     QΥ
                                                                                                                                                                                                                                                                            В
                                                                                                                                                          QΥ
                                                                                                                                                                                  DЬ
                                                                                                                                                                                                      δÃ
 Д
                   Qy
                                           밁
                                                               QY
                                                                                        Вþ
                                                                                                            Ωy
                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Ma
Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS;
PRINTS;
Repeat;
NON_TER
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                204
                                                                                                                                                                                    469
                                                                                                                                                                                                                                                                              372
                                                                                                                                                                                                                                                                                                                            316
                                                                                                                                                                                                                                                                                                                                                483
                                                                                                                                                                                                                                                                                                                                                                        260
                                                                                                                                                                                                                                                                                                                                                                                               429
                                                                                                                                                                                                                                                                                                                                                                                                                     200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                144
                                                                                                                                        528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        84 HNPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESS 143
                       864
                                                                                           574
                                                                                                                                                            707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local
 689
                                                                                                                 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                KPSYPPTYKPKPSYPPSYKTKKTYPSSYKAKPSYPPTYKAKPSYPPTYKAKPSYP----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETK 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSYPPTYGSKTNYLPLAKKLSS----YKPIKTTYNAK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLD
                                                                                                                                                                                                                                                                                                 APTTPKEPAPTTPKKPA--PTTPKEPA--PTTPKEPAPTTTKKPAPTAPKEPAPTTPKET
                                                                                                                                                                                                                                                                                                                           SYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAKPSYP---
                                                                                                                                                                                                                                                                                                                                               A--PTTTKEPS--PTTPKEPA-PTTTKSAPTTTKEPA-PTTTKSAPTTPKEPSPTTTKEP
                                                                                                                                                                                                                                                                                                                                                                      TYKAKPTYKAKPTYPSTYKAKPSYPPTYKAKPTYKAKPSYPPTYKAKPSYPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                TTTKEPA-PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPA--PTTPKKPA--PTTPKEPA- 428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEPTPTT----PKEPASTTPKEPTPTTIKSAPTTP----KEPAPTTTKSAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALITP
                                                                                                                                                                                                 PKEPAPTTPKEPAPTTPKE----PAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAP
                                                                                                                                                                                                                                              SYPPTYKAKPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKP-----T
                                                                                                             TTLKEPA--PTTPKKPA--PKELAPTTTKG-PTSTTSDKPAPTTPKETAPTTPKEPA--P
                                                                                                                                                            KELAPTTTKEPTSTTSDKPAPTTPKGTAPTT----PKEPAPTTPKEPAPTTPKGT---AP
                                                                                                                                                                                  PSYPASYKAKPSYPPTYKSKSSYPSSYKPKKTYPPTYKPKLTYKPTYK-PKPSYPPSYKP
                                                                                                                                        KTTYPPTYK-----
AKPTNPSTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKPTYPSTYKAKP-
                     ALENSPK----EPGVPTT--KTPA----ATKPEMTTTAKDKTTERDLRTTPETTTAAPK
                                             TYKAKPTYPSTYKAKPSYPPTYKPKISYPPTYKAKPSYPPT-YKAKPSYPPTYKAKPTYK
                                                                    TTPKKPA-PTT---
                                                                                          TYKAKPSYPPTYKAKPSYKAKPTYPSTYKAKPSYPPTYKAKPSYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 281;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PR01216;
PR01217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hydroxylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRICHEXTENSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADHESIVEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -YPPTYKAKPSYPSSYKPKKTYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----VKTSYSAPYKPPTYQPLKKKVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100412
                                                                    -PETPPPTTSEVSTPTTTKE----PTTIHKSPDESTPELSAEPTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26
                                                                                                                                                                                                                                   -PSYKAKPSYPPTYKAKPTYKAKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              . 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --KPKMTYPPTYKPKPSYPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MW.
                                                                                                                                        - PKISYPPTYKAKPSYPATYKAKPSYPPTYKAKPSYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 559; DB 1;
Pred. No. 1.9e-15;
2; Mismatches 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TANDEM REPEATS P-P-[ST].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6EA85312748CAACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OF Y-K-[PA]-K-[LP]-[ST]-Y-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     403;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----PKLTY----PPTYKP 143
                                                                                                                                                                                                                                     ------YPSTYKA----K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -TNYPPVY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -TYKSKPTYKPK
                                                                                                PTYKAKPSYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     260;
                                                                                                                                                                                                                                                                                                                                 -PTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -YRPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            75
                                                                                                                      812
                                                                                                                                                                                                                                                             652
                                                                                                                                                                                                                                                                                                           592
                                                                                                                                                                                                                                                                                                                                                                              315
                                                                                                                                                                                                                                                                                                                                                                                                  482
                                                                                                                                                                                                                                                                                                                                                                                                                         259
                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                               371
                                                                                                                                                                                                                                                                                                                                                       536
                                                                          863
                                                                                                                                                                                          527
                                                                                                                                                                                                                706
                                                                                                                                                                                                                                       468
                                                    688
                                                                                                 629
                                                                                                                                             573
                                                                                                                                                                   759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56
```

1-89

FROM N.A

```
THE THE TERM AND DESCRIPTION OF THE TERM AND DESCRIPTION O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; J05582; AAA60019.1; EMBL; M32738; AAA35806.1; EMBL; M32738; AAA35806.1; EMBL; M32739; AAA53806.1; EMBL; M52739; AAA59876.1; EMBL; K52229; CAA36478.1; AEMBL; X52228; CAA3647.1; AEMBL; M35093; AAB59612.1; AEMBL; M35093; AAB59672.1; EMBL; M31823; AAA578973.1; EMBL; M31823; AAA35757.1; EMBL; S81781; AAA14376.1; AEMBL; M21863; AAA14376.1; AEMBL; M21863; AAA14376.1; AEMBL; M31823; AAA14376.1; AAA14376; AAA143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M21868; AAA59874
PIR; A35175; A35175.
PIR; B35175; B35175.
PIR; S10218; S10218.
               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                               PROSITE; PS50024; SEA; 1
Glycoprotein; Signal:
                                                                                                                                                                                                                                                                                                                                                                                                                MIM; 113720,
InterPro; IPR000082; SEA
Pfam; PF01390; SEA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboratic between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on it use by non-profit institutions as long as its content is in no was modified and this statement is not removed. Usage by and for commercia entitles requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlycoSuiteDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <del>:</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Buluwela L., Liu Q., Lugmani Y.A., Gomm J.J., Coon Submitted (OCT-1992) to the EMBL/GenBank/DDBJ data -i- FUNCTION: DIRECT OR INDIRECT INTERACTION WITH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ru C.J., Yang P.C., Shew J.Y., Hong T.M., Yang Lee L.N., Luh K.T., Wu C.W.;
"Mucin mRNA expression in lung adenocarcinoma tissues.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Breast carcinoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Oncology 53:118-126(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF
TISSUE-Lung;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1-46 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=96181716; PubMed=8604237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALTERNATIVE PRODUCTS: VARIOUS VARIANTS ARE PRODUCED BY ALTERN SPLICING.
TISSUE SPECIFICITY: ABERRANTLY EXPRESSED IN HUMAN EPITHELIAL TUMORS, SUCH AS BREAST CANCER.
PTM: HIGHLY GLYCOSYLATED (N-AND O-LINKED CARBOHYDRATES AND S:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC. VARIES FROM 21 TO 125 IN THE NORTHERN EUROPEAN POPULATION. MOST FREQUENT ALLELES CONTAINS 41 AND 85 REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: CONTAINS 1 SEA DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CYTOSKELETON.
SUBCELLULAR LOCATION:
IS ALSO PRODUCED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACID)
                                                                                                                                                                                                                                                                                                                               Alternative splicing.
            24
1163
1187
1187
1034
957
975
1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P15941; -.
            23
1255
1162
1186
1255
1255
1255
1257
960
1151
975
1029
1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEA.
                                                                                                                                                                                                                                                                                                                                                        Cytoskeleton; Actin-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ALT_INIT.
ALT_INIT.
ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALT_SEQ.
ALT_SEQ.
ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE I MEMBRANE PROTEIN. A SECRETED
         N-LINKED
N-LINKED
                                                                                                                     EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
44 x 20 AA TANDEM REPEATS.
SEA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .
                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hong T.M., Yang
D (GLCNAC. . .)
D (GLCNAC. . .)
D (GLCNAC. . .)
D (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Coombes R.O databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cell lines and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S.C.,
) (POTENTIAL).
) (POTENTIAL).
) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACTIN
                                                                                                                                                                                                                                                                                                                                                        Transmembrane;
                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POPULATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BY ALTERNATIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Y.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIALIC
```

```
밁
                                                                         Qy
                                                                                                    밁
                                                                                                                           Š
                                                      B
                                                                                                                                                    8
                                                                                                                                                                          Ωy
                                                                                                                                                                                                     В
                                                                                                                                                                                                                           δõ
                                                                                                                                                                                                                                                   멅
                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                   B
                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 8.9
Best Local Similarity 27.8
Matches 295; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARSPLIC
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
VARSPLIC
      701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                    584
                                                                                                                                                                            697
                                                                            791
                                                                                                                           745
                                                                                                                                                                                                   473
                                                                                                                                                                                                                           645
                                                                                                                                                                                                                                                   417
                                                                                                                                                                                                                                                                                                    357
                                                                                                                                                                                                                                                                                                                          537
                                                                                                                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                           482
                                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                           441
                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      250 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16 TVLTV-----VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
APDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAP----DTRPAPGSTAP--
                     TPTTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEM
                                              DTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGS--TAPPAHGVTS
                                                                                                                     TPKEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTT
                                                                                                                                             GSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAPDTRPAPGSTAPPAHGVTS----
                                                                                               TRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAP
                                                                                                                                                                                            PPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP----PAHGVTSAPDTRPAP
                                                                                                                                                                                                                    PKA----AAPNTPKEPAPTTPK----EPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEP
                                                                                                                                                                                                                                               GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP---
                                                                                                                                                                                                                                                                    APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTT
                                                                                                                                                                                                                                                                                            GVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAH
                                                                                                                                                                                                                                                                                                                    ----APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEP
                                                                                                                                                                                                                                                                                                                                             SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT-SAPDT--RPAPGSTAPPAH
                                                                                                                                                                                                                                                                                                                                                                 PAPTTTKEPSPTTPKEPAPTT---TKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEP--
                                                                                                                                                                                                                                                                                                                                                                                            SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                            SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SAPTTPKEPAPTTP------KKPAPTTPKEP-----APTTPKEPTPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPGSTAPPAHGVT-SAPDT--RPAPGSTAPPAHGVTSAPDT--RPAPGSTAPPAHGVT- 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEP-----APTTPKEPAPTTTKEPAPTTTK 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P-KAETTTKGP--ALTTPKEP-----K 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        134
154
1021
1251
1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1255
2
134
154
1021
1251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1133
19
22
31
905
                                                                                                                                                                      KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8.8%; 5c.
27.8%; Pre
                                                                          -ETAPTTPKEPAPTTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87 FLQIYKGGGFL -> VSIGLSFPMLP (IN SECRETED ISOFORM).

55 MISSING (IN SECRETED ISOFORM).

7 T -> A (IN REF. 11).

84 P -> O (IN REF. 9).

94 P -> O (IN REF. 9).

122072 MW; 5E28DFC4DE7D9A82 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                   -APTTPKEPAPTAPK----KPAPTTPKEPAPTTPK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC. . .) (I
T-> TANTAPKDAT (IN ISOPO
MISSING (IN ISOPORM C).
MISSING (IN ISOFORM D).
MISSING (IN POLYMORPHIC I
ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 651; DB 1;
Pred. No. 7.1e-19;
2; Mismatches 460
                                                                      ----KKPAPTTPETPPPTTSEVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            460;
                                                                                                                                                                                                                                             PAHGVTSAPDTRPAPGSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ISOFORM B).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EPITHELIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          234;
                                                                                                                                               APD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                              700
                                                                      834
                                                                                               642
                                                                                                                     790
                                                                                                                                                                                                                                              472
                                                                                                                                                                                                                                                                                            416
                                                                                                                                                                                                                                                                                                                     584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           125
                                                                                                                                             583
                                                                                                                                                                                             528
                                                                                                                                                                                                                       696
                                                                                                                                                                                                                                                                     644
                                                                                                                                                                                                                                                                                                                                             356
                                                                                                                                                                                                                                                                                                                                                                   536
                                                                                                                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                            299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          51;
```

```
망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            рь
Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                         MUCL_HUMAN
STANDARW;
MUCL_HUMAN
MUCL_HIMAN
STANDARW;
MUCL_HIMAN
STANDARW;
MUCL_HIMAN

1.5941; p15942; p13931; p17626; Q14128; Q16444; vavv---
p15941; p15942; p13931; p17626; Q14128; Q16444; vavv---
p15941; p15942; p13931; p17626; Q14128; Q16444; vavv---
1.38820000 (Rel. 13, Created)
20-AUG-2001 (Rel. 40, Last annotation update)
MUCLN 1 PRECURSOR (POLYMORPHIC EPITHELIAL MUCLN) (PEM) (PEMT)
MUCLN 1 PRECURSOR (POLYMORPHIC EPITHELIAL MEMBRANE ANTIGEN) (EMA) (H23AG)
(TUMOR-ASSOCIATED EPITHELIAL MEMBRANE ANTIGEN) (EMA) (H23AG)
(TUMOR-ASSOCIATED EPITHELIAL MEMBRANE ANTIGEN) (EMA) (H23AG)
(TUMOR-ASSOCIATED ANTIGEN)
(TUMOR-ASSOCIATED ANTIGEN)
(TUMOR-ASSOCIATED ANTIGEN)
(TUMOR-ASSOCIATED ANTIGEN)
(TUMOR-ASSOCIATED ANTIGEN)
                                                                                                                                                                                                                                                                                                                                                                                                                         _HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1017
                               SEQUENCE FROM N.A. MEDLINE=90202794;
                                                                                                                                                                                                                    Homo sapiens (Human Eukaryota; Metazoa;
                                                                              Lan M.S., Batra S.K., Qi W.-N., Metzgar "Cloning and sequencing of a human pancy J. Biol. Chem. 265:15294-15299(1990).
                                                                                                                                                                                                        Mammalia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                612
                Ligtenberg M.J.L.,
                                                                                                                            MEDLINE=90368716;
                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              838
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       719
 Episialin,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPSSSTTESSSAPVT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTEKTTESKITAT---TTQVTSTTTQDTTPFKITTLKTTTLAPKVTT----TKKTITTT-EI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SAPVPTPSSSTTESSSAPVPTPS-SSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PYPTPSSSTTESSSAPYT----SSTTESSSAPYT----SSTTESSSAPYPTPSSSTTESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPIDTVFTRC---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SPKTVTTTVPCSTGTGEYTTEATTLVTTAVTTTVVTTESSTGTNSA-GKTTTGYTTKSVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              {\tt EGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSDETNICNGKPVDGLTTLRNG}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STSPSETA -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTSGCSPKTYTTTVPTTTTTSVTTSSTTTTTTVCSTGTNSAGETTSGCSPKTITTTVPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GT-TVTPSSSKYPGSQTETSVSSTTETTIVPTKTTTSVTTPSTTTITTTVCSTGTNSAGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PVPTPSSSTTESSSAPVSSSTTESSVAPVPTPSSSSNITSSAPSSI----PFSSTTESFST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PVPTPSSSTTESSSAPVPTPSSSTTESSVAPVPTPSSSSNITSSAPSSTPFSSSTESSSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PT-TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTTYLTTIAPTP - - SVTTVTNFTPTT - ITTTVCSTGT
                                                                                                                                                                                                        Eutheria;
                                                                                                                                                                                           s (Human).
Metazoa; Chordata; (
Metazoa; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -TTPKKPAPTTPETPPPTTSEVSTPT----TTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPT-TPKGTAPTTL
   carcinoma-associated
                                  PubMed=2318825;
                                                                                                                             PubMed=2394722;
                   Vos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -SESTTTSPTTPVTTVVSTTVVTTEYSTSTKPGGEITTTFVTKNI
                 H.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                - SSTTESSSAP-VPTPSSSTTESSSAPVPTPSSSTTESSSA
                   Gennissen
                                                                                                                                                                                                        Craniata; Vertebrata; Catarrhini; Hominidae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -AVPTTITT--TECSAATNAAGET
                                                                                           etzgar R.S.,
pancreatic
   ssen A.M.
mucin, i
                                                                                                                                                                                                          Hominidae;
    S.
                 C
   ., Hilkens
generated
                                                                                              Hollingsworth M. tumor mucin cDNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TKAPKKPTSTKKPKTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -NCEGKT
                                                                                                                                                                                                                                                                                       (H23AG) (PT
                                                                                                                                                                                                                          Euteleostomi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NSAGETTSGC
   J.;
by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -TTPKEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1207
                                                                                                 CDNA
                                                                                                                                                                                                                                                                                     ANTIGEN
                                                                                                                                                                                                                                                                                                       (PEANUT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  859
```

```
REFERENCE OF SERVICE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gene: an expressed Biochem. Biophys. [5] [5] SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J. Biol. [3]
Weiss M., Baruch A., Keydar I., Wreschner D.H.; Preoperative diagnosis of thyroid papillary carcinoma transcriptase polymerase chain reaction of the MUC1 gen Int. J. Cancer 66:55-59(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-90276413; PubMed=2351132;
MEDLINE-90276413; PubMed=2351132;
Transparent D.H., Hareuveni M., Tsarfaty
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gendler S.J., Lancaster C.A., Taylor-Papadimitriou J., Duhineat N., Burchell J., Pemberton L., Lalani E.-N., Wilson D. "Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin.";
J. Biol. Chem. 265:15286-15293(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
TISSUE=Breast carcinoma;
MEDLINE=90368715; PubMed=1697589;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91033045; PubMed=1688329; Jeltsch J.M., Hareuveni M., Horev J., Jettsch J.M., Garnier J.M., Lathe R., "Isolation and characterization of an coding for a breast-cancer-associated Gene 93:313-318(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J., Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschnie "A transcribed gene, containing a variable number of tandem codes for a human epithelial tumor antigen. cDNA cloning, exoff the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in breast cancer of the transfected gene and over-expression in the transfecte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-91097524;
Lancaster C.A., Po
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=Breast carcinoma;
MEDLINE=90276414; PubMed=2112460;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Human epithelial tumor antigen cDNA sequences. may generate multiple protein forms."; Eur. J. Biochem. 189:463-473(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lancaster C.A., Peat N., Duhig T.
Taylor-Papadimitriou J., Gendler
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        termini.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polymorphic gene encoding splice variants
                                                                                                                                                                                                                                                                                      MEDILIME=YOUNGER, T. Kufe D.;
Abe M., Siddiqui J., Kufe D.;
"Sequence analysis of the 5' region of the human
"Sequence analysis of the 5' region of the human
carcinoma-associated antigen gene.";
carcinoma-associated antigen gene.";
Res. Commun. 165:644-649(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Keydar
                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=90088473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  expressed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Burchell J.
"A highly i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gendler S.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PARTIAL SEQUENCE | MEDLINE=88330762;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A MEDLINE=91033045;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tissue.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [9]
                                                                                                                                                                                                                                                                  [10]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Biochem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             re and expression of the human expressed VNTR unit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chem.
                                                                                                                                                                                                                                    OF 1-109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  immunogenic region of a human by carcinomas is made up of ta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCE FROM N.A.
762; PubMed=3417635;
Taylor-Papadimitriou
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    265:5573-5578(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          rcinomas is made up of 263:12820-12823(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           189:475-486(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PubMed=2268309;
eat N., Duhig T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Res. Commun. 173:1019-1029(1990)
                                                                                                                                                                                                                                       FROM
                                                                                                                                                                                                                                                                                                                                                                                                                            PubMed=2597151;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FROM N.A.
                                                                                                                                                                                                                                       N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       and over-expression in breast cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wilson
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zaretsky J., Weiss M.,
, Keydar I., Wreschner D.H
n expressed hypervariable
d antigen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tandem
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polymorphic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Duhig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          м.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polymorphic epithelial mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Smorodinsky N. Lathe R., Dion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Τ.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            repeats
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Differential splicing
                                                                                                                                                                                                                                                                                                                                                                          DF3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rothbard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         epithelial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wreschner
                                                                             oma by
gene.'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Duhig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            expression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A.,
                                                                                                              reverse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          D.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Þ
```

```
RESULT

ANCHOL

ANCH

ANCHOL

ANCHOL

ANCHOL

ANCHOL

ANCHOL

ANCHOL

ANCHOL

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                       STRAIN=S288C / AB972;
Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V., Walsh S.V., Whitehead S.;
Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AMYH_YEAST STANDARD; PRT; 1367 AA. P08640; P08068; 01-AUG-1988 (Rel. 08, Created) 01-FEB-1995 (Rel. 31, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update) GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAM GLUCOSIDASE) (1.4-ALPHA-P-GLUCAM GLUCOHYDROLASE). STAL OR STA2 OR MAL5 OR YIRO19C.
                                                                                                                            SEQUENCE OF 1-31 FROM N.A. STRAIN-SPX101-1C;
MEDLINE-89031230; pubMed=3141213; Pardo J.M., Ianez E., Zalacain M. "Sinilar short elements in the 5' from Saccharomyces cerevisiae."; FEBS Lett. 239:179-184(1988).
                                                                                                                                                                                                                                                          "Gene
                                                                                                                                                                                                                                                                       SEQUENCE OF 1-242 AND 762-1331 FROM MEDLINE-87194600; PubMed-3106330; Yamashita I., Nakamura M., Fukui S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Fungi; / Saccharomycetales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Saccharomyces cerevisiae (Baker's yeast)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                810
                                                                                                                                                                                                        Bacteriol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YIKRSESV----ALINRALERGPLNGAPKLFPDVNESYWAF-GDIMD
                                                                                                                                                                                                                                                        fusion is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PARRITEVWGIPSPIDTVFTRCNCEGKTFFFKD--SQYWRFTNDIKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIDISNPKFDD----CVGHWAQEFIEKLTSLGYISGYPDGT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---IINPMLSDETNICNG----KPVDGLTTL-----RNGTLVAFRGHYFWMLSPFSPPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LADTHWAAWAIKFATSQGLFKGYPDGTFKPDQNITRAEFATVVLHFLTKVKGQEIMSKLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YLRGYPDGSFRPERNITRAEAAVIF-----AKLL----GADESYGAQSASPYSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LLRPH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGGSGGSGGGGGGGTVPTSPTPTPTSKPTSTPAP---TEIEEPTPSDVPGAIGGEHRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSD----EPTPSDEPTPSDEPTPSDEPTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAPTTPKKPAPT-TPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----TSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHM-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATNSKAT -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEPTPSDEPTPSETPEEPTP----TTTPTPTPSTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKEPGVPTTKTPAAT--KPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTPS--DEPTPSETPEEPIPTDTPSDEPTPSDEPT----PSDEPTP--SDEPTPSDEPTP
                                                                                                                                                                                                   169:2142-2149(1987)
                                                                                                                                                                                                                                                  akamura M.,
a possible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----TPKPQKPTKAP-KKPTSTKKPKTMPRVRKPKTTPTPRKM-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ascomycota: Saccharomycotina: Saccharomycetes: Saccharomycetaceae; Saccharomyces
                                                                                                                                                                                                                                                  mechanism
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        --VEMPE-----VTPDMDYLPRVPNQGI-----
                                                                                                                                                                                                                                                                       S
                                               , Claros
regions
                                                                                                                                                                                                                                             underlying
                                               of M
                                               f the
                                            Jimenez
e STA2 ar
                                                                                                                                                                                                                                             the evolution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1,4-ALPHA-
                                            ez A
and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1201
                                          SGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FKPQN
                                                                                                                                                                                                                                             of.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -PTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                X >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               926
```

```
Ъ
                                                                                                Qy
                                                                                                                                              밁
                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                 Дb
                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SO THE TENT OF THE PROPERTY OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
Matches 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                entities
or send a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collab between the Swiss Institute of Bioinformatics and the EMBL outst the European Bioinformatics Institute. There are no restrictions use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for contition and the statement is not removed.
                  581
                                                         560
                                                                                                                                            513
                                                                                                                                                                                        462
                                                                                                                                                                                                                                    464
                                                                                                                                                                                                                                                                           405
                                                                                                                                                                                                                                                                                                                      408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Signal; Multigene SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                358
                                                                                                                                                                                                                                                                                                                                                                                                         354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL 1,4-LINKED GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUCING ENDS C WITH RELEASE OF BETA-D-GLUCOSE.

-!- SIMILARITY: TO S.POMBE SPBC215.13.
-!- SIMILARITY: SOME, TO S.POMBE SPC285.13C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3L; Z38061; CAA86176.1; -.
3L; M16164; AAA35014.1; -.
M16165; AAA35015.1; -.
3L; X13857; CAA32069.1; -.
3; B26877; B26877.
4; A26877; A26877.
5; S48478; S48478.
PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPT-TPEEPTPTTPEE
                                                                                                                                                          PTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPT----TTPKEPAPTTKEPAPTTPKEPAPT
                                                                                                                                                                                                                                                                                                                                           APTT----TKSAP----TTPKEPAPTTKEPAPTTPKEPAP----TTTKEPAPTTTKSA
                                                                                                                                                                                                                                                                                                             APVTSSTTESSSAPVTSSTTESSSAPVTSS---
                                                                                                                                                                                                                                                                                                                                                                                                         APV----
                                                                                                                                                                                                                                                                                                                                                                                                                                   APTSKYLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEP
                                                                                  PTTPKEPSPTTTKEPAPTTPKEPAPT-TPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KKKTKKVIESEEITEEHSVSENOESSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DLSTGCNNYDNQGHSQTDFPGFYWNIDCDNNCGGTKSSTTTSSTSES--STT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DCDAQCKKYDK----CCPDYESFCAEVHNPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPN 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PT--PSSSTTESSS-----APVPTPSSSTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTTTSCTKEKPT - - PPHHD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S0001458; MUC1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -TSSTSESSTTTSSTSESSTTTSSTSESSTSSTT-----APATPTTTSCTKEKPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       requires a license agreement (S
an email to license@isb~sib.ch).
                                                                                                                                     -PTPSSSTTESSSAPAPTPS---SSTTESSSAPVTSS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      210
817
874
1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Glycosidase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                         PTPSSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       family.
21
1367
1367
1367
817
874
                                                                                                                                                                                                                   -SSTTESSSAPV----PTPSSSTTESSSAPVT---SSTTESSSAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    136110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polysaccharide degradation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLUCOAMYLASE S1/S2.
SER/THR-RICH.
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          809; DB 1;
No. 6.le-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                         -TTESSSAPVTSSTTESSSAPVPTPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----ESSSAPVTSSTTESSS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                -TTESSSAPVPTPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF THE CHAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                commercial
    639
                                              611
                                                                                                                                559
                                                                                                                                                                          521
                                                                                                                                                                                                                   512
                                                                                                                                                                                                                                                             461
                                                                                                                                                                                                                                                                                                         463
                                                                                                                                                                                                                                                                                                                                                  404
                                                                                                                                                                                                                                                                                                                                                                                         407
                                                                                                                                                                                                                                                                                                                                                                                                                                     357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1ts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53;
```

```
γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               рЬ
                                                                                                                                                                                                                                                                                                                                                                                                               망
                                                                                                                                                                                                                                                                                                                                                          멍
                                                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                    SLP1_CLOTM
SLP1_CLOTM
Q06852;
Q06852;
Q1-JUN-1994 (Rel. 29, Created)
Q1-JUN-1994 (Rel. 29, Last sequence update)
Q1-FEB-1996 (Rel. 33, Last annotation updat)
Q1-FEB-1996 (Rel. 33, Last annotation updat)
Q1-FEB-1996 (Rel. 33, Last annotation updat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000
STRAIN-NCIE 10682;
MEDLINE-93209931; PubMed-8458832;
Fujino T., Beguin P., Aubert J.-P.;
Fujino T., Beguin P., Aubert thermocellum gene cluster encoding the "Organization of a Clostridium thermocellum gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in attachment of the cellulosome to the cell surface.";
J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          581
                                                                                                                                                                                                                                                                                                                                                             2350
                                                                                                                                                                                                                                                                                                                                                                                       1050
                                                                                                                                                                                                                                                                                                                                                                                                                   2290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      598
                                                                                                     SEQUENCE FROM N.A.
                                                                                                                             NCBI_TaxID=1515;
                                                                                                                                            Bacteria; F:
Clostridium
                                                                                                                                                                        Clostridium thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QTPTTTPITTTTVTPTPTPTGTQTPTTTPITTTTTTTPTPTPTGTQTPTTTPITTTTTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEP----APTTPKEPAPT-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TQPTMTTTTTENPTPPTTTPITTTTTTTPTFTFTGTQTPTTTPITTTTTTTTPTPTPTFTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPT-TPEEPAPTTPKAAAPNTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NISCRATMYPDVPIGQLGQTVVCDVSVGLICKNEDQKPGGVIPMAFCLNYEINVQCCECV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PLPPSITPPTFSPFSTTTPTTPCVPLCNWTGWLDSGKPNFHKPGGDTELIGDVCGPGWAA 1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKEPAPTTP----KETAPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTTTPSSPI-TTTTPSSTTTPSPPPTTMTTPSPPTTTMTTLPPTTTSSPLTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKELAPTTTKGPTSTTSDKPAP-TTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EPTSTTSDKPAP-TTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P-TGTQTPTTTP--ITT--TTTVTPTPTPTGTQTPTTTPITTTTTTTTTPTPTPTGTQTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QVTSTTTQDTTPFKITTLKTTTLAPKVTTT-KKTITTTEIMNKPEETAKPKDRATNSKAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TKPEMTTTAKDKTTERDLR----TTP--ETTTAAPKMT---KETATTTEKTTESKITATTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTPTGTQTPTT---TPITTTTTTTPTPTPTPTGTQTPTTTTTTTTTTTPTPTPTGTQTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                   TPITTTTTVTPTPTGTQTPTTTPITTTTTTTTTPTFTGTQTPTTTPITTTTTTTVTPTFT
                                                                                                                                                                                                                                                                                                                                                                                                                                           TPKPQKPTKAP-KKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNP---TSRIAEAML
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEP---
                                                                                                                                                                                                                                                                                                                                                            PTGTQTPTTTTTTTTTTTPTPTPTG----TQTP--TTTPITTTTTTTP
                                                                                                                                                                                                                                                                                                                                                                                    QTTTR-PNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTA
                                                                                                                                                           Firmicutes;
                                                                                                                                                              Bacillus/Clostridium
                                                                                                                                                                                                                    update)
                                                                                                                                                              group;
                                                                                                                                                                                                                     LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----GVPT-TKTPAA
                                                                                                                                                               Clostridiaceae;
                                                                                                                                                                                                                        PROTEIN
                                                                                                                                                                                                                                                                                                                                                                  2392
                                                                                                                                                                                                                                                                                                                                                                                              1096
                                                                                                                                                                                                                        ₿)
                                                                                                                                                                                                                       (S-LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   934
```

```
Qy
                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                          рь
                                                                                                                                                                                                                                                                                                                                                                20
                                                                                                                                                                                                                                                                                                                                                                                          B
                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                            g
                                                                                                                                                                  Qy
                                                                                                                                                                                           DЪ
 밁
                  QY
                                             В
                                                                Ωy
                                                                                            뮹
                                                                                                                    Qy
                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIGNAL CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            \frac{1}{1} \cdot \frac{1}{1} \cdot \frac{1}{1}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00395;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             entities requires a license agreement (S or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROSITE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
 1209
                                                                                                                                              1040
                                                                                                                                                                                                                      530
                                                                                                                                                                                                                                            920
                                                                                                                                                                                                                                                                                              865
                                                                                                                                                                                                                                                                                                                                                                    362
                                                                                                                                                                                                                                                                                                                                                                                           758
                                                                       695
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBDUIT: ASSEMBLED INTO MONO-LATERED CRISTALINE PROPERTION: CELL WALL.
SUBCELLULAR LOCATION: CELL WALL.
SINILARITY: CONTAINS 4 S-LAYER HOMOLOGY (SLH) DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                 VLAKPTP-KAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTT 361
                                                                                                                                                                                                                                                                                                                                                                 TKSAPTTPKEPAPTTTKEPAPTTPKEPA----PTTTKEPAPTTTKSAPTTPKEPAPTTPK 417
                                                                                                                                                                                                                                                                                                                                                                                           VVIQPAPIKAASDEPIPTDTPSDEPTPS-----DEPTPS----DEPTPSDEPTPSD
                                                                                                                                                                                                                                                                                            SDEPTPSDEPTPS--DEPTPS--DEPTP-SDEPTPSETPEEPIPTDTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                 KPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPT-TPKEPAPTAPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                           EPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTP
                                                                                                                                           TPSDEPTPSE----TPEEPIPTDTPSDEPTPSDEPTPSDEPTPSD-EPTPSDEPTPSETP
                                                                                                                                                                   APT---TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPT---TPEELAPTTPEEPTPT-TP
                                                                                                                                                                                             PSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEP
                                                                                                                                                                                                                   ----PTTTKEPAPT-TPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEP
                                                                                                                                                                                                                                            SDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPT
                                                                                                                                                                                                                                                         X67506; CAA47841.1;
                                                                                              EEPIPTDT9SDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSD
                                                                                                            EEPAPTTPKAAAPNTPKEPAPTTPKEPA--PTTPKEPAPT-TPKETAPTTFKGTAPTTLK
---DEPTPSDEPTPSETPEEPI-
                       PKGTAPTTLKEPAPT-TPKKPAPKELAPTTTKGPTSTTSDKPAPT---TPKETAPTTPKE
                                                            PS01072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR001119; SLH
                                                                                                                                                                                                                                                                                                                                                                                                                                          12.8%;
nilarity 31.0%;
Conservative 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        s-layer;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1378
1453
1495
1566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29
36
36
207
409
607
771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ASSEMBLED INTO MONO-LAYERED CRYSTALLINE ARRAYS
                                               -DEPTPSD-EPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPS--DEPTPS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SLH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A
A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1664
763
191
363
565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SLH_DOMAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1449
1494
1565
1625
1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         763
1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       178194
                                                                                                                                                                                                                                                                                                                                                                                                                                             104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPROXIMATE TANDEM REF

T-P-S-D-E-P.

GLY/PRO/SER/THR-RICH.

SLH 1 (INCOMPLETE).

SLH 2.

SLH 3.

SLH 4 (INCOMPLETE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
CELL SURFACE GLYCOPROTEIN
4 X 156 AA APPROXIMATE REF
                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 (INCOMPLETE).
5F396695BA9FE74B
                                                                                                                                                                                                                                                                                                                                                                                                                                                        950;
   PTDTPSDEPTPSDEPTPSD-EPTPSDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1;
.6e-30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1664
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                            476
                                                                                                                                                                                                                                                                                                                                                                                                   804
                                                                                                                                                                                                                                                                           529
                                                                                                                                                                                                                                                                                                     919
                                                                                                                                                                                                     1039
                                                                                                                                                                                                                            584
                                                                                                                                                                                                                                                     979
                                                                                                                             694
                                                                                                                                                     1094
                                                                                                                                                                              637
                                                      1208
                                                                              753
                                                                                                                                                                                                                                                                                                                                                                                                                                                    55
```

```
THE PROPERTY OF THE PROPERTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRINTS; PRO0438; GFCYSKNOT.
SMART; SM00214; VWC; 2.
SMART; SM00211; VWC; def; 2.
SMART; SM00215; VWD; 4.
SMART; SM00216; VWD; 4.
PROSITE; PS00122; EGF_1; UNKNOWN_1.
PROSITE; PS01125; CTCK_1; 1.
PROSITE; PS01125; CTCK_2; 1.
PROSITE; PS011208; VWFC; 2.
                    REPEAT
DOMAIN
DOMAIN
DOMAIN
DOMAIN
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000359; Cys_knot.
InterPro; IPR000561; EGF-11ke.
InterPro; IPR002400; GF Cysknot.
InterPro; IPR001007; VWFC.
InterPro; IPR001007; VWFC.
InterPro; IPR001846; Vwd.
Pfam; PF00007; Cys_knot; 1.
Pfam; PF00094; vwd; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; L21998; AAB95295.1; EMBL; M74027; AAA59875.1; EMBL; M94131; AAA59163.1; EMBL; M94132; AAA59164.1; MIM; 158370; ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and content is in no waverentities requires a license agreement is no usage by and content is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                    163
423
670
770
790
1139
1154
1215
1230
1246
1787
1820
                                                                                                                                                                                              Repeat;
                  Signal.
              POTENTIAL.
MUCIN 2.
APPROXIMATE REPEATS
          (POTENTIAL)
                                                                 밁
                                                                                            Q
                                                                                                                           밁
                                                                                                                                                   õ
                                                                                                                                                                                   ₽
                                                                                                                                                                                                           δÃ
                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                      δõ
                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                Db
                                                                                                                                                                                                                                                                                                                                                                                          δõ
                                                                                                                                                                                                                                                                                                                                                                                                                        B
                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
Matches 403
         1641
                                                                                                                           1521
                                                                                                                                                                                                                                             1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                     464
                                                                                                                                                                                   1461
                                                                                                                                                                                                                                                                                                      1362
                                                                                                                                                                                                                                                                                                                                                              1302
                                                                                                                                                                                                                                                                                                                                                                                                                         1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                            196
                                                                                                                                                                                                                                                                                                                                                                                                                                                    163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19
PSPPPTTTTPPPTTTPSPPTTTTPSPPITTTTPPPTTTPSSPITTTPSPPTTMTTPS
                          KKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAP-
                                                                       PAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEP----
                                                                                                               TTPSPPTTTTTPPPTTTTSPPTTTPTTPTSTTTLPPTTTTPSPPTTTTPPPTTTTPPPTTTPSP
                                                        PTTTTPSPPTITTTPPPTTTPSPPTTTTTPPPTTTPSPTTTPITPPTSTTTLPPTTT
```

```
TTTPPPTTTPSPPTTTPSPPTTTPSPPTTTTTPPPTTTTPSPMTTPITPPASTTTLppT
                                                                                        TTKGPALTTPKEPT-----PTTPKEPASTTPKEPTPTTIKSAP-TTPKEP------APT
                                                                                                                                                                                        TTKSAP-TTPKEPAPTTTKEPAPTTPKEP-----APTTTKEPAPTTTKSAP--TTPKE
                                                                                                                                                                                                                                                   GFICKNEDQFGNGPFGLCYDYKIRVNCCWPMDKCITTPSPPTTTPSPPPTTTTLPPTTT
                                                                                                                                                                                                                                                                                                                                            CCLWSDWINEDHPSSGSDDGDREPFDGVCGAPEDIECRSVKDPHLSLEQHGQKVQCDVSV 1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NSAANRELOK------KLKVKDNKKNRTKKKPTPKPPVV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ESFCAEVHNPTSPPSSKKAPPPSGASQTI----KSTTKRSPK--PPNKKKTKKVIESEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -CGPNGTVEKHENICSITTRPSTLTTFTTITLPTTPTSFTTTTTTTTTPTSSTVLSTTPKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CGCYVEDTHYPPGASVPTEETCKSCVCTNSSQVVCRPEEGKILNQTQDGAFCYWEI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E--CEWHYEPCGNRSFETCRTINGIHSNISVSYLEGCYPRCPKDRPIYEEDLKKCVTADK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SCSCDTGGDCECFCSAVASYAGECTKEGAC-----VFWRTPDL-CPIFCDYYN---PPH
                                                                                                                                                                                                                                                                                                                                                                                                    ------DEAGSGLDNGDFK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---ITEEH---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               bal Similarity
403; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                  ----ITTAKPINPRPSLPPNSDTSKETSLT
                                                                                                                                                          -TTTTPPPTTTPSPPITTTTTP------LPTT----TPSPPIST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9 4391
2 4362
3 4373
3 4427
8 4438
8 4438
8 4438
6 4616
6 4616
7 44627
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
7 4787
9 5019
9 5069
9 5069
9 1449
9 1449
9 1449
9 1449
9 1449
9 1449
9 1449
9 1449
9 1449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -SVSENQESSSSSSSSSSSSSTI------WKIKSSK 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16
29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . 5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Score 1188; DB 1;
; Pred. No. 4.2e-39;
85; Mismatches 536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      > L (IN REF. 3).
> S (IN REF. 3).
> P (IN REF. 3).
> T (IN REF. 3).
> S (IN REF. 3).
> S (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                             -VTTPDTSTTQH-NKVSTSPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 5179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                      410
                                                                                                                                                          1460
                                                                                                                 360
                                                                                                                                                                                                       313
                                                                                                                                                                                                                                                     1421
                                                                                                                                                                                                                                                                                                  253
                                                                                                                                                                                                                                                                                                                                                                                                                                            1301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53;
```

1580

1640 463

```
Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                        Minimum
Maximum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Run
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                Result
No.
                                                                                                                                                                                                                                                                                                                                                        Database
                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        on:
pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                        DB DB
1188
8090
8090
8050
8551
5551
5551
5517.5
509.5
509.5
509.5
509.5
509.5
4498.5
4498.5
4498.5
4498.5
4498.5
4498.5
4498.5
4498.5
4498.5
4498.5
                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                                                                                                                                                                                                                                                        seq
                                                                                                                                                                                                                                                                                                                                                                                                        length:
                                                                                                                                                                                                                                                                                 Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-556-246-1_COPY_25_1404
7410
1 QDLSSCAGRCGEGYSRDATC...
                                                                                                                                                                                                                                                                                          Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           April 26, 2002, 16:19:03; Search time 47.4 Seconds (without alignments) 1067.457 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
 SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                100059 seqs, 36664827 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QDLSSCAGRCGEGYSRDATC.....ARAITTRSGQTLSKVWYNCP 1380
                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                         0
20000000000
 DΒ
MUC2_HUMAN
SLP1_CLOTM
AMYH_YEAST
MUC1_HUMAN
FP1_MYTED
ZAN_HUMAN
EXTN_TOBAC
NFH_MOUSE
TCNA_TRYCR
CPN_DROME
FP1_MYTCO
NFH_HUMAN
NFH_HUMAN
NFH_HUMAN
NFH_HUMAN
NFB1_MOUSE
BAT2_PLAYO
EXTN_MAIZE
ZAN_MOUSE
HWP1_CANAL
VRP1_YEAST
VGLX_HSVEB
YJ9P_YEAST
YAV1_SCHPO
VG50_HUMAN
MAPA_BAT
T2D1_MYTGA
T2D1_MYTGA
T2D1_HUMAN
TCOF_HUMAN
TCOF_HUMAN
TCOF_HUMAN
                                                                                                                                                                                                                                                                                   IJ
                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                    100059
 Q05049 xenopus lae P24928 homo sapien P12036 homo sapien P08775 mus musculu P11414 cricetulus P48634 homo sapien Q01443 plasmodium P14918 zea mays (m O88799 mus musculu P46593 candida alb P37370 saccharomyc P29968 equine herp P47179 saccharomyc Q10172 schizosacch Q00130 ictalurid h Q9y2n3 homo sapien P34926 rattus norv Q27409 mytilus gal Q00268 homo sapien P46013 homo sapien
                                                                                                                                                                                          Q02817 homo sapien
Q06852 clostridium
P08640 saccharomyc
P15941 h mucin 1 p
Q25460 mytiucin 2 p
Q25460 mytiucin sedu
Q9y493 homo sapien
P13983 nicotiana t
P19246 mus musculu
P23253 trypanosoma
Q02910 drosophila
                                                                                                                                                                            Q02910
Q25434
P16884
                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                    mytilus cor
                                                                                                                                                                          rattus norv
```

	O O O O T T T T T T T T T T T T T T T T	PRESENTATION OF THE PRESEN	RES
i			29 2222000000
	SEQUENCE OF 626-1895 TISSUE-(Colon; Pul Gum J.R. Jr., Hicks , Kim Y.S.; "The human MUC2 inte both upstream and do J.B. Jr., Hicks . SEQUENCE OF 1343-189 SEQUENCE OF 1343-189 MEDLINE-91358717; MEDLINE-91358717; MICC2 human small i and polymorphism."; J. Clin. Invest. 88: -i- FUNCTION: COATS OTHER MUCUS MEMB PROTECTIVE, LUBR	77; 1 V-1 V-1 V-1 V-1 Sap Yot Lia Tax Tax NCE E=I I Cullifition	5 433.5 4 433.5 6 428.6 7 427.5 8 427.5 9 425.5 10 424.5 11 424.5 12 423.5 13 422.5 14 419.5 14 411.5 15 411.5 17 1
AGENTS AT MUG SUBUNIT: MULI SUBUNIT: MULI PISSUE SPECII BRONCHUS, CEI BRONCHUS, CEI POLYMORPHISM POLYMORPHISM VARIES AMONG SIMILARITY: USIMILARITY:	CE OF 626-189: =COLON; E=93016075; PR R. Jr., Hicks R. Jr., Sum Para N. W., Gum Para N. W., Gum Para N. W., Gum Para N. W., Kim Para N. W., Kim Para N. W., Kim Para N. W., Sum Pa	(Rel. (Rel. (Rel. (Rel. CURSOR C. C. (Hum Metazo utheri 9606; 9606; 9606; 100 N.A sstine; 32002; 1100 of	5.77 5.77 6.77
CONTAL SOCIAL SO	AN J.W Sti Sti Sti 137 137 137 137 137 1100	29, 35, 40, (INT n). Ch ; Pr ; Pr ; Pr ; Pr ; Pr ; Pr ; Pr ; Pr	3164 339 3 2517 3 1251 7 1185 7 1127 7 3027 7 3421 7 247 7 247 8 907 6 907
ORFI ON: S COI OESI OESI OESI OESI OESI OESI OESI OE	ND 419 W: Too W:	The section of the first terms o	
ETED. SMALL INTESTINE, CO SMALL INTESTINE, CO BLADDER. ARE INVOLVED IN INT (BY SIMILARITY). (BY SIMILARITY). F REPEATS IS HIGHLY ELES. DOMAIN SHOWS SOME S C DOMAINS. C DOMAINS. ERMINAL CYSTINE KNOT	96-5179 FROM N.A. 100449; 10149; 1016-1419 In as cysteine-rich serit repetiti 1383 (1992). 176-4195 FROM N.A. 185763; Culhane P.J., Lagace R. 13(1991). 13(1991). 13(1991). THELIA OF THE INTESTINES, THELIA OF THE INTESTINES, THALINING ORGANS. THOUGHT BARRIER AGAINST PARTICLE	uence update) otation update) MUCIN 2). Craniata; Vertebrata; Catarrhini; Hominidae; Catarrhini; Hominidae; catarrhini Hominidae;	TEGU_HSV11 XP2_XENLA NCR2_HUMAN NCR2_HUMAN TRX2_HUMAN DRPL_HUMAN MAP4_MOUSE SGS3_DROME TEGU_HSVEB DRPL_RAT ZAN_PIG VGP3_EBV ALIGNMENTS PRT: 5179 AA.
LONIC TUMORS RACHAIN OR POLYMORPHIC SIMILARITY TO	Laga bdoma e reg e reg e reg o PRO	Homo. Homo. Kim Y. DNA.	p10220 p17437 Q9y618 Q9umn6 Q9umn6 P274259 p27840 p02840 p028955 p54255 p54255 p63200
ORS, R R TO THAT	ce R.E., ins located ion."; cks J.W., cks J.W., ed arrays NTDE A INFECTIOUS	stomi; S.; imilarity	herpes simp xenopus lae h nuclear r caenorhabdi homo sapien mus musculu drosophila equine herp rattus norv sus scrofa epstein-bar

\$ 				1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	E.	e i ge			Č.				7	*	÷	1			e de					7			. u. , , ,	3- 4-
																												•
																									jac.			
		ri di	•. •	N. T.						- 12			·							ě			Y.					
		a l					9		-/1		3						ii i	in the second				7		, , , , , , , , , , , , , , , , , , ,		1:		
			a a													- 52°	# 55°		in a second	3.54 2.54	3. ()		5				*	
				<i>3</i> *=																						4	8	And the second s
								ķ.				T.																And the second s
	3.				ž.																							

```
extensin class 1 precursor - cowpea
C; Species: Vigna unguiculata (cowpea)
C; Species: Vigna unguiculata (cowpea)
C; Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Nov-2000
C; Accession: T11622; S54155
R; Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
Mol. Plant Microbe Interact. 10, 95-101, 1997
A; Title: Rhizobia modulate root-hair-specific expression of extensin genes.
A; Reference number: Z17301; MUID:97155574
A; Accession: T11622
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-489 < ARS)
A; Cross-references: EMBL:X91836; NID:g1015936; PIDN:CAA62943.1;
A; Cross-references: EMBL:X91836; NID:g1015936; PIDN:CAA62943.1;
A; Cross-references: EMBL:X91836; NID:g1015936; PIDN:CAA62943.1;
A; Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
Submitted to the EMBL Data Library, April 1995
A; Description: A class of root-hair specific extensins involved
A; Reference number: 554155
A; Accession: S54155
A; Status: preliminary
A; Status: preliminary
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: Ext26G
C;Superfamily: hydroxyproline-rich glycoprotein
C;Keywords: glycoprotein; hydroxyproline
C;Keywords: glycoprotein; hydroxyproline
F;1-23/Domain: signal sequence #status predicted <NAT>
F;24-489/Product: extensin class 1 #status predicted <NAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: mRNA
A;Residues: 326-489 <AR2>
A;Cross-references: EMBL:X86030; NID:g791149; PID:g791150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Genetics:
                                                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                                                 Dр
                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                               δÃ
                                                                                                                                                                                                        Ъ
                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                          밁
                                              ΔÃ
                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Sim
Matches 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      471
                                                                                                                                                                                                                                                                                                                                                                                                              188 YYYKSPPPPSPSPPPYYYKSPPPPSPSPPPPYYYKSPPPPSPSPPPYYYKSPPPPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           354 PKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAP 413
                                                                                                                                                                                                                                                              651
                                                                                                                                                                                                                                                                                                               248
                                                                                                                                                                                                                                                                                                                                                                591
                                                                                                      367
                                                                                                                                                                                                             308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKSAPTTTKSAPTTPKEPSP 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PSPSPPPPYVYKSPPPPSPSPPPPYYYKSPPPPS--PSPPPPYYYKSPPPPSPSPPPPYY 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTP---KEPAPTAPKKPAPTT 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PKQTPPYYYNAPPYYYKSPPP<sup>c</sup>----PSPSPP------PPPYYHKYPPYYYKSPPP 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPK 590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YKSPPPPSPSPPPPYYYKSPPPPSPSPPPPYYYKSPP--
                                                   KKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPP 828
                                                                                                                                                                                                           PSPSPPPPYYYKSPPPPSPSPPPPYYYKSPPPPSPSPPPPYYYYKSPPPPSPSPPPPYYY-
                                                                                                                                                                                                                                                         NTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELA 710
                                                                                                                                                                                                                                                                                                                  ETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAP
                                                                                                      ----RSPPPPSPYYYYYYQQQQSASAYYYYYQQQQQSAYYYYYQQQQSXYYYYY
                                                                                                                                                     PTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP 770
SPP----PPYVYKSPPPPSPSPPPYYYKSPPPPSPSPPPYYYKSPPPPSPSPPPPP 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.5%; Score 633; DB 2; Length 489; 31.2%; Pred. No. 1.4e-19; tive 33; Mismatches 252; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----PPSPSPPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PID:g1015937
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              in rhizobium/legume inte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                          307
                                                                                                                                                                                                                                                                                                                                                                             650
                                                                                                                                                                                                                                                                                                                                                                                                                           247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7;
```

Search completed: April 26, 2002, 16:18:41 Job time: 311 sec

40;

```
A; Gene: xNopp18
C; Superfamily:
C; Keywords: pho
                                                                                                                                nucleolar phosphoprotein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
C;Accession: I51618; S57757
R;Cairns, C.; McStay, B.
J. Cell Sci. 108, 3339-3347, 1995
A;Title: Identification and cDNA cloning of a Xenopus nucleolar
A;Reference number: I51618; MUID:96019267
A;Accession: I51618
                                                                           G
                                                                                                                                                                                                                                               RESULT
I51618
                                                                                                A:Status: preliminary; translated A:Molecule type: mRNA A:Residues: 1-990 <CAI>
                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                 A; Cross-references: EMBL: X88927; NID: g895920; PIDN: CAA61368.1;
                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                               ş
                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                문
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                    Superfamily: nucleolus-cytoplasm; Reywords: phosphoprotein
   Query Match
                                                                                                                                                                                                                                                                                              2938 TSTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                          2772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2258
                                                                                                                                                                                                                                                                                                                                                                   997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         622
                                                                                                                                                                                                                                                         14
                                                             xNopp180
                                                                                                                                                                                                                                                                                                                                                                                                    TKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTT-----------
                                                                                                                                                                                                                                                                                                                    LOTT
                                                                                                                                                                                                                                                                                                                                        PPPTTTTTAKPATTSGKRGPPSIQPPAEMFTTPAP-----PPPSNGGYGEETNQEEEQV
                                                                                                                                                                                                                                                                                                                                                               PQKPTKAPKKPTSTKKPKTMPRVRKPK---TTPTPRKMTSTMPELN-----PTSRIAEAM
                                                                                                                                                                                                                                                                                                                                                                                      SSTPRVRLITGTPDDLIVSVTVPSHGNRRQNITASSV---PSNSTSPIILPSES-LTTPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                     R----APAEITMSSESHREISTVSSEPSEPEIPLSTTVSPNVVTASSIPSEEPILSSVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEPETNAPAVTVSSEASSTTLEENSSTSSP----TSSEASVKLSSLFPESITSEAVTVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KSPDESTP--ELSAEPTPKALE--NSPKEPGVPTTKTPAATK----PEMTTTAKDKTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ESTTESSEAPTTPAKTSETKPSNVSSTSRKSTENVETSTSQSGSLESSTMSSTS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -STTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTSPKSLVKSTTSPSTVTSSEPSESTKRTTVSTTVSTTTPTEETTTSESLILTAAPSKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VPKESESTTTSSESSKPSQEPAGILTSTVVVPTSSVSLITASEIEAITSNTPFKQGRTPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EFSKVPESTISSVLEEDLTKTTPSPILEETTTASETSEPLTEDSLTVSVRIHELTTSSEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPDTTTASEEPTKSTSMSPDLSTTSNVLSESSTTPESSS----KSPVSSSTEGISVVTST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NSEKPT-SAPELVTSSVTHVASSSPDVPTES--SEPDDLTGSSTENIPEASSKQTISSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                             -TTPK-----EPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APKELAPTTTKEPTSTTSDKPA------PT------TP--KGTAP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EELAPTTPEEPTPTTPEEP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TKKPAPTAPK-----EPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TVE----TPKTSSEVSLNSEEPSTTEAPTTLSPDILSTTTNNLSQSSTVSTEDRSEISSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --PKEPAPT-----TPKEPAP-----TTPKETAPTTPKGTAPTTLKEPAPTTPKKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----STEQPTSTTPSGQSLTPMNSNSEVLTTSEPHVLSSSLSPDVSQSSTTPNNLSESS
                                                                                                                                                                                                                                                                                             2942
                                                                                                                                                                                                                                                                                                                    1053
   8.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----APTTPKAAAPNTPKEPAPTT------
                                                                                                                       from GB/EMBL/DDBJ
   Score
                                                shuttle
  643.5;
                                               phosphoprotein
  BG
  2
                                                                                                                                                       nucleolar phosphoprotein, xNopp18(
 Length 990
                                                                                  PID:g895921
                                                                                                                                                                                                       24 - Sep - 1999
                                                                                                                                                                                                                                                                                                                                                                1048
                                                                                                                                                                                                                                                                                                                                        2937
                                                                                                                                                                                                                                                                                                                                                                                       2883
                                                                                                                                                                                                                                                                                                                                                                                                                                      2827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            897
                                                                                                                                                                                                                                                                                                                                                                                                              996
                                                                                                                                                                                                                                                                                                                                                                                                                                                             942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    621
                                                                                                     8
                                                                                                                                B
                                                                                                                                                    δÃ
                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                            Db
RESULT
                                                                                 B
                                                                                                                                                                              밁
                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                ğ
                                                                                                                                                                                                                                                                                                                                                                         멍
                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                             ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 24.2 Matches 246; Conservative
                                                                                  844
                                                                                                                                                       877
                                                                                                                                                                                                     817
                                                                                                                                                                                                                                                                            644
                                                                                                                                788
                                                                                                                                                                              758
                                                                                                                                                                                                                             704
                                                                                                                                                                                                                                                    757
                                                                                                                                                                                                                                                                                                  704
                                                                                                                                                                                                                                                                                                                          591
                                                                                                                                                                                                                                                                                                                                                  644
                                                                                                                                                                                                                                                                                                                                                                          558
                                                                                                                                                                                                                                                                                                                                                                                                 584
                                                                                                                                                                                                                                                                                                                                                                                                                        509
                                                                                                                                                                                                                                                                                                                                                                                                                                                525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               57
15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88
                                                                                                                                                                              QAA-
```

```
886 EAVTPE----
                                                     991 KATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTP-RKMTSTMPELNPTSRIAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             302 -- KVLAKPTPKAETTTKGPALTTPKEPTPTTPKE--PASTTPKEPTPTTIKSAPTTPKEP
                                                                                                                                                             TSTTTQDTTPFKITTLKTTTLAPKVTTTK--KTITTTEIMNKPEETA-KPKDRATNS--- 990
                                                                                                                                                                                                                                                                             TKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQV 936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPK 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PAPKELAPTTTKEPTSTTSD-----KPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAKPAVKTT---PGKATSK---PVVASKPVP-AKKASSSSDSDSSEEETTKTTKPLTKLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKK
                                                                                                                                                                                                                        SSDEDVSKAKKTNTAVSKSPV----TTPKAVPAAKKESSSESSDSEDEKQGGKNTSTTKI 843
                                                                                                                                                                                                                                                                                                                                                                                           KPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPT 876
                                                                                                                                                                                                                                                                                                                                                                                                                                                   SSEEEGKSKQP-TGKSPAAKATAP-PKKNPVAVNKDKPSSSSSSDSSGDDEKQ----KPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KKDSSSSDSSSDEKKTPA----KRAAKTTPAKPAAKTTPAKPA-----AKTTPAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEP-APTTPKEPAPTTTKKPAPTAPKE 583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPA-PTTPKEPAPTAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DSSSEDEKKSSVKLGVKAAP--KKAPA-----APDAKSTPVAAAKKSAPAKKASSSSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APTT-----TKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKE-PAPTTTKSAP 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KKKTLSQPGTKAKPESSDSSDSSDEEEQPAKKAKIVPAKAAASAPKPLAKKAETSTDSES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPTPKPP---VVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPN 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SPDAKKRKPPANGLPKKKSAKESSSSEDSSSEEDEPPAKKRAQPAGGKKPVVKAVQPKKA 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAVKTLPPKKAESSSDSSSSSSSSKKTKPAKPPAKSATPVNTKAPAQNKASKASCSDSDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SSSDSSDSSSDEETTTKPAAKTTPAKSAATPTSKTPTNSKATPTSKKTPAKPGTPKTSAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SDSSSNEETTTKPAAKTTPAKSA-ATPTSKTPTNGK-ATPTSKTPAKPGTPKTSTAKKDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --TKPATAAKPQAKKTAGKKSSSREDSSDSSDEEQKTAKSKPKPDVYSAVP---PPTSVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QPVIKVPPKQAVVKAGLASNNG----KTADSSSSEDSDSPPAKKTAATK-----TPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KSSSEDSSDESDSEEETKKPPAKRPAQTPKVAAVKTPTQKKAKSSSSESSSSEDEASKKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSVSENQESSSSSSSSSSSTIWK-IKSSKNSAANRELOKKLKVKDNKKNRT-----KK 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPPSSKKAPPPSG------ASQTIKSTTKRSPKPPNKKKT-----KKVIESEEITEE 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SSSSDSSDSSEDEKKS
-NKKLKAKSPNTFPKVNKKELKNTPFRRVVEEDIEINP--RMAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24.2%; Pred. No. 1.1e-19; ative 148; Mismatches 424;
                                                                                                          PKAAAAECSEESSSEDEGKANGTSGKRKRESTGNAEC 885
                                                                                                                                                                                                                                                                                                                                         ----AAKDVKQGAKAAKPTPKKAASSSSE---DS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               199;
                                                        1046
  932
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   357
                                                                                                                                                                                                                                                                                                                                         787
                                                                                                                                                                                                                                                                                                                                                                                                                                                   757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      224
```

1133	1075 -GAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSDETNICNGKPVDGLTTL	Qу
3217	3158 TPTATMSTATPSSTPETVHTSTVLTTTATTTGATGSVATPSSTPGTAHTTKVPTTTTTGF	DЬ
1074	EDAG-	Qy
3157	3098 STTATTPTVTSSKATPSSSPGTATALPALRSTATTPTATSFTAIPSSSLGTTWTRLSQTT	Db
1032		Qy
3097	:: : :	Db -
866		Qy
3037	TPSSTPGTTWILTELTTAATTTAGTGFTATPSSTPGTTWIL	Db
946		Qy
2977	2918 SPRTATTLPVLTSTATKSTATSVTPIPSSTLGTTGTLPEQTTTPVATMSTIHPSSTPETT	Db
907	PETT	Qy
2917	2861 PGTTWILTELTTTATTTASTGSTATP-SSTPGTAPPPKVLTSPATTPTATSSKATSSS	Ф
878	1	VQ
2860	2801 LGELGQVVECSLDFGLVCRNREQVGKFKMCFNYEIRVFCCNYGHCPSTPATSSTAMPSST	Db
826	811APTTPKKPAPTTPETP	Qy
2800	2741 VTTGCEPQCAWSEWLDYSYPMPGPSGGDFDTYSNIRAAGGAVCEQPLGLECRATAQPGVP	Db
810	811	Qy
2740	LLPSOPTSAPITTV	Db
810	KELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEP	Qy
2684	ATTGTTTTSTPALS	рь
763	- I'K	Qy
2624	STVTPSSALGTTHTPPVPNT	Db
729	681 APTTPKGTAPTTLKEPAPTTPKKPAPKELAPT-TTKEPTSTTSDKPAPTT	Qy
2566	TTTPTTSGSTVTPSSIPGTTHTPTVLTTTTQPVATGSMATPSSSTQTSGTPPSLIT	Db
680	TPKEPAPTTPKET	Qy
2506	TTTTGFTVTPSSSPGTARTPPV	Db
632	603PEELAPTTPEEP 6	Qγ
2446	PMATMSTATPSSTPE	Db
602	566 KEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTP- 6	Qy
2386	2333 GTTWILTKLTTTATTTESTGSTATPSSTQGPPAGTPHVSTTATTPTVTSSKATP	Дb
565	508 TTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTP 5	Qy
2332	2273 RELGQVVECSLDFGLVCRNREQVGKFKMCFNYEIRVFCCNYGHCPSTPATSSTATPSSTP 2	DЪ
507	- 'D	Qy
2272	13 TTGCEPQCAWSEWLDYSYPMPGPSGGDFDTYSNIRAAGGAVCEQPLGLECRAQAQPGVPL	Db
491	485 TTTKEPS	Ov
2212		Db
484	PKEPAP	Qy

```
R; Favello, A.; Vaudin, M.
submitted to the EMEL Data Library, August 1994
A; Description: The sequence of C. elegans cosmid
A; Reference number: 221536
A; Accession: T34513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hypothetical protein ZK783.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T34513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ОÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Genetics:
A; Gene: CESP: ZK783.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA A; A; Molecule type: DNA A; Residues: 1-3507 <FAV> A; Residues: 1-3507 <FAV> A; Cross-references: EMBL:Ul3646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1 A; Experimental source: strain Bristol N2; clone ZK783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ър
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Map position: 3
A;Introns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 586/1; 3504/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                         멍
                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                       DЬ
                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   рь
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QУ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                       ŠÕ
                                                                     DЪ
                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 286; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3276 ATGSM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1134 RNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPIDTVFTRCNCEGKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1781 MGSCGCKCMAGYTGDGATCIKIEEEPKSDKTA----CTDEWSRLCELE---KKQCTVDE 1832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1886 PDGFIGDGMICDDVDECNNAGMCDDENTKCENTIGSFNCVCLEGFKKVDEKCVVDEKKQP 1945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1833 EEVPQCGACLPGHHPINGTCQSLQISGLCAQKNDCNKHAE--CIDIH-----PDSHFCSC 1885
                                                                                                                                                                                                                                                 2080 TVKLSSKSPEVTESSVKSSPSTPS-TTSQSVTSTVPETSKSTVLSSEAPVTSTSPTEVHT
                                                                                                                                                                                                                                                                                                                                                                                                                                1971 GIVSSTS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1946 NREKIEIDEENSSSSNSG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           132 EEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPK 191
                                                                                                                                                                                                                                                                                    304 LAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKS-APTTPKEPAPTTT
                                                                                                                                                                                                                                                                                                                                                                                 249 ETSLTVNKETTV-ETKETTTINKQTSTDGKEKTTSAKETQSIEKTSAKDLAP----TSKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              55 ERGREC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 LSSCAGRCGEGYSRD-ATC-----NCDYNCOHYMECCPDFKRVCTAELSCKGRCFESF 54 : | | : | : | : | : | : | : |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TATPSSSPGTALTPPWISTTTTPTTTTP--TTSGSTVTPSSIPGTTHTARVLTTTTTV 3275
EPAPTTTKEPSPTTP--KEPAPTTTKSAPTTTKEPAPTTTKSAP-----
                                                                                                                                                                                                   KSAPTTPKEPAPTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPT 422
                                                                                                                                                                                                                                                                                                                                       EVSTSSSKSTTASETTVSSTPSESSSSEAPLTSSPATTTEVITESSVKSTTPKEESSSEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPN---SDTSK 248
                                                                   TKTFDATESSTVQASETSSGTSVKSTSEPESHVTKLSITSSNPSSSVPVTSPKSTPTVPE 2257
                                                                                                                                                           -SSETKPSLSASSTTGDTNSTTPSTSSLASVKSTSAPEGTSASVAPVKLSSLSPDVSQPS
                                                                                                          TPK-EPAPTTPKEPTPTTPKEPAPTTKEP-APTTPKEPAPTAPKKPAPTTPKEPAPTTPK 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative 178; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                  ----ATSSESTTAEPHVTTSISSTTSTKDMTSSKSPENVTMSSESP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8.9%; Score 658.5; DB 2; 22.6%; Pred. No. 8.7e-20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -ATPSSS----TOTSGTPPSLTTTATTITATGST 3309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -DCDAQCKKYDKCCPDYESFCAEVHNPTSPPSSKKAPP 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -ASQTIKSTTKRSPKPPNKKKTKKVIESEEIT 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  484;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZK783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  605/1; 1175/3; 1207/1; 1409/
           -TTPKEPSPTT 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3507;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     317;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QEKPTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                   2079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1970
                                                                                                                                                                                                                                                             2138
                                                                                                                                                                                                                                                                                                                                                                                                                                         2019
                                                                                                                                                                                                                                                                                                       362
                                                                                                                                                                                                                                                                                                                                                                                             303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51;
```

```
밁
                                                                                                                                                      Š
      덩
                           S
                                                                                              Ş
                                                                                                                            B
                                                                                                                                                                                       밁
                                                                                                                                                                                                                  ş
                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                         문
                                                                                                                                                                                                                                                                                                                                                                                                     ş
                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local
      7710
                                                                                                                                                                                                                                                    7491
                                                                                                                            7598
                                                                                                                                                                                       7541
                                                                                                                                                                                                                                                                                                               7438
                                                                                                                                                                                                                                                                                                                                                                         7380
                                                                                                                                                                                                                                                                                                                                                                                                                                     7324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     7108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6996 LVPKKEAVPPAKGRTVLEEKVSVAFRQEVVVKERLELEVVEAEVEEIPEEEEFHEVEEYF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7056
                                                                                                                                                                                                                      761
                                                                                                                                                                                                                                                                                  701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        74 PDYESFCAEVHNPTSPPSSKKAPPPSGASQTIKSTTKRS-----PKPPNKK 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
nes 268; Conserv
      PKPEA--
                                                                                                                      NPQVPEKVELTPLKVPGGEKKVRKLLPERKPEPKEEVVLKSVLRKRPEEEEPKVE----P
                                                                                                                                                                                                             TLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETA-----PTTPKEPAPT
                                                                                                                                                                                                                                            YEE -- PEEIAP---EEEIAPEEEKPVPVAEE----EEPEVPPPAVPEEPKKIIPEKKVP-
                                                                                                                                                                                                                                                                                                                                     NTPKEPAP----TTPKEP-----APTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                             LT----PTTPEKLAPTT----PEKPAPTTPEELAPTTPEEPAPTTPEEPAPTTPKAAAP
                                                                                                                                                                                                                                                                                                                                                                                                                               TTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNK--
                                                                                       KALENSPKEPGVPTTKTPAATKP----EMTTTAKDKTTERDLRTTPETTTAAPKMTKETA
                                                                                                                                                     TPKKP--APTTPETPPPTTSEVS--TPTTTKEP-----TTIHKSPDESTPELSAEPTP
                                                                                                                                                                                  VIKKPEAPPPKEPEPEKV---IEKPKLKPRPPPPPPPAPPKEDVKEKIFQLKAIPKKKVPE
                                                                                                                                                                                                                                                                         PKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPT
                                                                                                                                                                                                                                                                                                        EVPKEVVPEKKVAVPKKPEVPPAKVPEVPKKPV----LEEKPAVP----VPERAESPPPEV
                                                                                                                                                                                                                                                                                                                                                                    VTVPPKNPVPEKKAPAVVAKKPELPPVKVPEVPKEVVPEKKVPLVVPKK--PEAPPAKVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKK 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EEGEFHEVEEFIKLEOHRVEEEHRVEKVHRVIEVFEAEEVEVFEKPKAPP------KG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KTKKVI--ESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK 177
                                                            KKLE-KVKKPAVP---EPPPPKPVEEVEVPTVTKRERKIPEPTKVPEIKPAIPLPAPEPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KEVVPEKKVPAAP--PKKPEVTPVK-VPEAPKEVVPEKKVPVP-PPKKPEVPPTK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPTTPKEPAPTTVKEPAPTTP-----KEPAPTAPKKPAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EVVP-----EKKVPVPPAKKPEAP----PPKVPEAPKEVVPEKKVPVPPPKKPEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKK----PAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PEISEKIIPPKKPPTKVVPRKEPPAKVPEVPKKIVVEEKVRVPEEPRVPPTKVPEVLPPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PALT----TPKEPTPTTP-KEPASTTPKEPTPTTIKSAPTTPKEP-APTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EVIKKAVPEAPTPVPKKVEAPPAKVSKKIPEEKVPVPVQKKEAPPAKVPEVPKKVPEKKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DNKKNRTKKKPTPKPPVVDEAGSGLDN--GDFKVTTP-DTSTTQHNKVSTSPKITTAKP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PPTKVPEVPKAAVPEKKVPEAIPPKPESPPPEVFEEPEESPSAPPK-KPEVPPVRVPEVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -INPRPSLPP---NSDTSKETSLTVNKETTVETK---ETTTTNKQTSTDGKE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PEEEEFVPE---EEVLPEVKPKVPVPAPVPEIKKKVTEKKVVIPKKEEAPPAKVPEVPKKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EEKRIILPKEEEVLPVEVTEEPEEEPISEE------EIPEEPPSIEEVEEVAPPRVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----SAKDLAPTSKVLAKPTPKAETTTKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.0%; Score 666.5; DB 2;
24.2%; Pred. No. 9.1e-20;
tive 120; Mismatches 491;
---EVKTIKPPPVEPEPTPIAAPVTVPVVGKKAE 7745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7962;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                     862
                                                                                                                                                                                                                 813
                                                                                                                                                                                                                                                                                                                                                                                                 650
                                                                                         918
                                                                                                                                                                                                                                              7540
                                                                                                                                                                                                                                                                                                        7490
                                                                                                                                                                                                                                                                                                                                                                                                                                7379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              370
                                                                                                                                                                                                                                                                                                                                       700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50;
```

```
밁
                                            δÃ
                                                                                                                        20
                                                                                                                                                                  Вþ
                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                           Ъ
                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 рЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Experimental source: placenta C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-3570 < DES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R;Desseyn, J.L.; Guyonnet-Duperat, V.; Porchet, N.; Aubert, J.P.; Laine, A. J. Biol. Chem. 272, 3168-3178, 1997
A;Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various A;Reference number: Z22899; MUID:97166151
A;Accession: T45025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mucin MUC5B, tracheobronchial [imported] - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Gene: MUC5B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Cross-references: EMBL: 272496; NID: g1834502; PIDN: CAA96577.1; PID: g1834503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 8.9%; Score 663; DB 2; Best Local Similarity 22.1%; Pred. No. 5.8e-20; Matches 396; Conservative 119; Mismatches 578;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Accession: T45025
        2093
                                                                                                                                                                  1973
                                                                                                                                                                                                                                                                                                                                                                                                           1794
                                                                                                                                                                                                                                                1913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1620
                                                                                                                                                                                                                                                                                                                             1853
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1581 VECSLDFGLVC------RNRE----QVGKF-KMCFNYEIRVFCCNYGHCP
                                                                                                                                                                                                                                                                                     277
                                                                                                                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                                                                                                                               201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30 MECCPDFKRVCTAELSCKGRCFESFERGRECDCDAQCKKYDKCCPDYE--SFCAEV-HNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12
                                                                                                                                                         LTPPVWISTTTPTTRGSTVTPSSIPGTTHTATVLTTTTTVATGSMATPSSSTQTSGTP
                                                                                                                                                                                                                                     PSSTPETVHTSTVLTTTATTTRTGSVATPSSTPGTAHTTKVPTTTTTGFTATPSSSPGTA
                                                                                                                                                                                                                                                                                                                                                         NPR------PSLPPNSDTSKETSLTVNKETTVET----KETTTNKQTSTDG 276
                                                                                                                                                                                                                                                                                                                                                                                                                                       GLDNG------KVSTSPKITTAKPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STPATSSTATPSSTPGTTWILTEQTTAATTTATTGSTAIPSSTPGTAPPP-----KVLT
PNTTATTHGRSLPPSSPHTVPTAWTSATSGILGTTHITEPSTGTSHTPAATTGTTQPSTP
                                                                                                               TTIKSAPTTPKEPAPTTTKSAP--TTPKEPAPTTT-KEPAPT----TPKEPAPTTTKEPA 397
                                                                                                                                                                                                     LTTP----KEPTPT-----TPKE-----------
                                                                                                                                                                                                                                                                                                                   SSRATPSSSPGTATALPALRSTATTPTATSVTAIPSSSLGTAWTRLSQTTTPTATMSTAT 1912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PMATMSTIHPSSTPETTHTSTVLTTKATTTRATSSMSTPSSTPGTTWILTELTTAATTTA 1793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSPPSSKKAPPPS----------GASQTIKSTTKRSPKPPNKKKTKKVIE 126
                                     PTTTKSA-----PTTP-----PKEP 427
                                                                                PSLTTTATTITATGSTTNPSSTPGTTPIPPVLTTTATTPAATSSTVTPSSALGTTHTPPV
                                                                                                                                                                                                                                                                           ALPHGTPSSTPGTTWILTEPSTTATVTVPTGSTATASSTRATAGTLKVLTS-TATTPTVI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SQATTPTATSSKATSSSSPRTATTLPVLTSTATKSTATSFTPIPSSTLGTTGTSQNRPPH 1733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEEITEEHSVSENQESSS------SSSSSSSSSTIWKIKSS----KNSAANR--- 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEIKDILTESEFVGSSAIFECLVSPST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---RKMTSTMPELNPTSRIAEAMLQTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AKAPKEEAAKPKGPIKGVPKKTPSPIEAERRKLRPGSGGEKPPDEAPFTYQLKAVPLKFV 7805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----PEETAKPKDRATNSKATTPKPQKPTKAPKKPTS--TKKPKTMPRVRKPKTTPTP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----ELQKKLKVKDNKKNRTKKKPTPKPP-----VVDE-----AGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1053
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3570;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels 698;
                                                                                                                                                                                                 --- PASTTPKEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-Jul-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
  2152
                                                                                                                                                                                                 344
                                                                             2092
                                                                                                                                                           2032
                                                                                                                                                                                                                                           1972
                                                                                                                                                                                                                                                                                                                                                                                                     1852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1619
                                                                                                                                                                                                                                                                                                                                                                                                                                         234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alter
```

```
A;Gene: GDB:MUC1; PUM
A;Gene: GDB:MUC1; PUM
A;Gene: GDB:MUC1; PUM
A;Gene: GDB:MUC1; Q23
A;notrons: 20/1; 62/3; 1165/3; 1184/2; 1230/1; 1270/3; 1320/3
A;notrons: 20/1; 62/3; 1165/3; 1184/2; 1230/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 20/1; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: mRNA
A; Residues: 1-142, '0', 144-162, '0', 164-168 < ABE>
A; Cross-references: EMBL: M31823; NID: g181542; PIDN: AAA35757.1;
R; Masuzawa, Y: Myauchi, T: Hamanoue, M.; Ando, S.; Yoshida,
J. Biochem. 112, 609-615, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Note: undetermined tyrosine residues in the carboxyl-terminal non-repetitive region C; Comment: This protein is length polymorphic. Individuals may have between 21 and 12 partial repeats. The repeat shown is defined by SmaI nuclease sites. C; Comment: Serine and threonine residues in the tandem repeat domain are extensively C; Comment: For an alternative splice form without a tandem repeat domain, see PIR:548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: mRNA
A; Residues: 998-1011, 'ES', 1014-1017;1018-1032, 'T', 1034-1037;1038-1057
A; Experimental source: gastric carcinoma cell
R; Zrihan-Licht, S.; Baruch, A.; Elroy-Stein, O.; Keydar, I.; Wreschne: FEBS Lett. 356, 130-136, 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Biochem. Biophys. Res. Commun. 165, 644-649, A;Title: Sequence analysis of the 5' region c A;Reference number: A36735; MUID:90088473 A;Accession: A36735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Title: A novel core protein as well as polymorphic epithelial mucin carry A;Reference number: JX0235; MUID:93123189
A;Accession: PX0066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Title: Tyrosine phosphorylation of the MUC1 breast cancer membrane proteins cytokine A; Reference number: 851026; MUID: 95080414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Contents: annotation
                                                                                                                                    Дb
                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                            QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QY
밁
                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                       135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  252 LTVNKETTVETKETTTTNK--QTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT
                                                                                                                                                                                                                                                                                       191
                                                                                                                                                                                                                                                                                                                                                           382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-KAETTTKGP--ALTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LTVLTATTAPKPATVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                    PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP-----
{\tt TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV}
                                                                       KPAPTTPKEPAPTTPK----EPAPTTTKEPSPTTPKEPAPTT---TKSAPTTTKEPAPTT
                                                                                                                                               TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                       PAPGSTAPPAHGVTSAPDT--RPAPGSTAPPAHGVTSAPDT--RPAPGSTAPPAHGVTSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTPTT----IKSAPTTPKEPAPTTT-----KSAPTTPKEPAPTTTKEP-----A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK 134
                                                                                                                                                                                                                                                                                           PDT--RPAPGSTAPPAHGVT-SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              310;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 672; DB 1;
Pred. No. 9.6e-21;
8; Mismatches 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -TPTTPKEPASTTP-----KE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              472;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        human DF3
                                                                                                                                                                                                                                                                                                                                                                      ----KKPAPTTPKEP----
                                                                                                                                                                                                                       -APTTPKEPAPTAPK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 1344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wreschner, D.H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           breast carcinoma-associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PID:g181543
J.; Takao,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
           367
                                                                                                                                                                                                                           464
                                                                                                                                                                                                                                                                                                                                                                      427
                                                                                                                                                                                                                                                                                                                                                                                                                                              190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      74
                                                                                                                                                                                                                                                                                                  247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            s.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              see PIR:S4814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    peanut
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Shimazu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 region a
and 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                91
```

```
Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                       C;Accession: 138346
R;Labeit, S.; Kolmerr, B.
Science 270, 293-296, 1995
A;Title: Titins: giant proteins in charge
A;Reference number: A57430; MUID:96026330
A;Accession: 138346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Species: Homo sapiens (man)
C; Date: 29-May-1998 #sequence_revision
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
I38346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           뭥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B
A;Cross-references: GDB:127867;
A;Map position: 2q31-2q31
                                                                                                                                                             A; Molecule type: mF
A; Residues: 1-7962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  elastic titin -
                                                                 A; Gene: GDB: TTN
                                                                                                      C; Genetics:
                                                                                                                                     A;Cross-references:
                                                                                                                                                                                                                                  A;Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TKSAPTTPKEPSPTTTKEP-----APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             T-SAPDT--RPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TLKTTTLAPKVTTTKKTITTT--EIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRPAPGSTAPPAHGVTSAP----DTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAPGSTAPPAHGVTSAPDTRPAPGS--TAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               {\tt APGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTR}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AP-PAHGVTSAPDTRPAPGSTAPPAHGVTS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APTITKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTPKEPAPTTPKG----TAPTTLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HGVTSAPDTRPAPGSTAP----PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGST 600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---EPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---ELAPTTPEEPTPTTP------EEPAP--TTPKA----AAPNTPKEPAPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---EPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ASGSASGSASTLVHNGTSARATTTPASKSTPFSIPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSDETN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STKKPKIMPRVRKPKTTPTPRKMTSTMPELNP----TSRIAEAMLQT-TTRP----NQTPN 1060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAP---GSTAPPAHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        {\tt TRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPTAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPPAHGVTSAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----TKPEMTTTAKDK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VTSAPDTRP---APGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  human (fragment)
                                                                                                                                                                                                       mRNA
                                                                                                                                                                           ARESY
                                                                                                                                     EMBL: X90569; NID: g1017426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTP--FKIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KKPAPKELAPTTTKGPTSTTSDKPAP -- TTPK ---- ETAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -KKPAPTTPETPPPTTSEVSTPTTTKEPTT--
                                         OMIM: 188840
                                                                                                                                                                                                                                            from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----GVTSAPDNRPALGSTA----PPVHNVTS
                                                                                                                                                                                                                                            GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29-May-1998 #text_change 21-Jul-2000
                                                                                                                                                                                                                                                                                                                                         of muscle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --APDTRPAPGSTAPPAHGVTSAPDTRP
                                                                                                                                                PIDN:CAA62189.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1085
                                                                                                                                                                                                                                                                                                                                            ultrastructure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KKPAPKEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -IHKSPD
                                                                                                                                            PID:g1017427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTERD
                                                                                                                                                                                                                                                                                                                                            and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        709
                                                                                                                                                                                                                                                                                                                                            elasticity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 772
```

```
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
945 TPFKITTLKTTTLAPKVT---TTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ---
                                          888 SSQKSEENESSAEKPGARRDFVPKKHKTTVKPAETTSAVAASTTTTEPITTTEKSTTLET 947
                                                                                                      910
                                                                                                                                                                                                                                                811 APTTPKKPAPTTPE--TPPPTTSE-VSTPTTTKEPTTIHKSPDESTPELSAEPTPKALEN 867
                                                                                                                                                                                                                                       798
                                                                                                                                                                                                                                                                                                                              738
                                                                                                                                                                                                                                                                                                                                                                        771
                                                                                                                                                                                                                                                                                                                                                                                                                  089
                                                                                                                                                                                                                                                                                                                                                                                                                                                              730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        523 EESSTTEEVTTTEEPÄNTGNPPTTENPTTTEQPTSTAESTTTALPFTTEQTVTTEEPTTA 582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  566 KEPAP----TTTKKPAPTA--PKEPAPTT---PKETAPTTPKKLTPTTPEKLA---PTTP 613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              464 FDEPTEIT-KAPHPGKLLEKKTYHFVLSDNFARYSEAKENDDYNHLDYNHYREAKEPTTT 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        543 --EPAPTTPKKPAP----------TTPKE------PAPTTP 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  404 KESDGFHTTLKLKVTTADSDSTESATTVKPFNEETTTKSHVVPKPTKKGTVKVTPKLELS 463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       345 VAKSEDKMTLSKTAATETTQQTTEVT-DGPEKETTKNVSIEIPITTVPLVETTSTTSTAS 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       398 PTTTKSAPTTPKEPAPTTP----KKPAPTTPKEPAPTTPKEP-TPTTPKEPAPTTKEPAP 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200 TEIASPMETNTTTEATTTSVEPSVSTLAS---EDETTVTAIAEST-----TTVIAEVST 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            141 PEITTGIVVINSKSESVTDMSTTRFSTTLSPTTELLTSPETLVSTDSSTST-EQTSPDNT 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             226 PKITTA-KPINPRPSLPPNSDT-----SKETSLTVNKETTVETKETTTTNKQTSTDGK 277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             169 ELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQ---HNKVSTS 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      109 TKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSTIWKIKSSKNSAANR 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ω
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             49
                                                                                                                                -----TVDTSSATTEESSTAAETTTTSAE---TSETTTSESAAFITGESPENTALQS 887
                                                                                                                                                                             SPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTA----------
                                                                                                                                                                                                                       TTTTTTTEKTTSKTTTEKPTTSESATTETTTSEPST----TEST----- 838
                                                                                                                                                                                                                                                                                                             VLKEKKRLLKEKESTSTTGSDSSETTTVVAENIDEVTTTEKEKVVQTTPITTEKSTTQEE
                                                                                                                                                                                                                                                                                                                                                       ----KKPAPKELAPTTIKGP------TSTTSDKPAPTTPKETAPTTPKEP 810
                                                                                                                                                                                                                                                                                                                                                                                                    DEETTSTTSTTPEITSTKE--IVTESAITQTSVSVVESSTPRQLPERWKAIVNKFKHNLE
                                                                                                                                                                                                                                                                                                                                                                                                                                              STTGKATTPELSTTSEETTTTELKITTE-----GSTTTEEPTTTAIFAEASTGIITT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAP----TT 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EKSTATQ----KPTTTQESVST---EKTSTTKKA---STTEE--PTTTDEPTTTT---ES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPA 673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKE-----PSPTTPKEPAPTTTKSAP 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA 397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .....INNDNVDIGPSGDSNPTGSSWFQEIEATVGGQTVKSEHNIDSSV 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------TITKEPAPT----TTKSAPTTPKEPSPTTTKE----PAPT-----TPK---- 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STTTSSETSTEK---PTTPLIDNKIAGPATGK---PETTHEPVTGTTPN--FDTATETPF 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RDFKIVKHIKKNTCTCSC----KCVPD------APSN----PFDVSTTISS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RCFESFERGRECDCDAQCKKYDKCCPDYESFCAEVHNPTSPPSSKKAPPPSGASQTIKST 108
                                                                                    ------APKMTKETATTTEKTTESKITATTTQ-VTST---TTQDT 944
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----TTSTDASTTNAPTTGKDSTT
                                                                                                                                                                             909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    70
```

Ş g S В Š B 5 밁 Ş

Š 밁

В Š 밁 Š В ş 멍 ş 밁 Š 밁 ρy В Ş В Ş 8 δ 밁 Š В Ş

A; Cross-references: EMBL:X52229; NID:g37053 R; Wreschner, D.H. Submitted to the EMBL Data Library, March 1990 A; Reference number: \$40293 A; Accession: \$40293 A; Molecule type: mRNA A; Residues: 1-19.29-155, 'P',157-175, 'P',177-182, 'A',184-212,1033-1037,' A; Cross-references: EMBL:X52229; NID:g37053; PIDN:CAA36478.1; PID:g37058 R; Abe, M.; Siddiqui, J.; Kufe, D.	69; PIDN:AAA59876.1; PID:g188 one copy of the tandemly repgar, R.S.; Hollingsworth, M./an pancreatic tumor mucin cDf 6 ranslation A',1341-1344 <lan> 98; PIDN:AAA60019.1; PID:g188 four fewer copies of the tancty, I.; Smorodinsky, N.; Hore cDNA sequences. Differential</lan>	WID:g182121; PIDN:AAA35804 PISIA2 present only the a WID:g182126; PIDN:AAA35806 PISIB2 present only the a Padimitriou, J.; Duhig, T Phuman tumor-associated	35175 35175 ucin 1 precursor, repetitive splice form A [validated] - human validated precursor, and precursor, and precursor an	Db 948 TPIEATTLNEVTGAPVDETTINTLELLSKINTQISQPKPTDIS Qy 999KPTKAPKKPTSTKKPKTMPRVRKPKTTDTPRKMTSTMPELN Db 998 KTDALSSLISGLIGSFTKAPMAPTIHTTTDAAFVTATEASLNDGSDK Qy 1040PTSRIAEA 1047 Db 1045 KIIDEAQPTDEIRRA 1059
1344 <wre> 1344 <wre> 1037,'A',1039-1344 <</wre></wre>	g188870 repeated sequence M.A. CDNA. cDNA: g189599 tandemly repeated se Horev, J.; Zaretsky, ial splicing may gen	4.1; PID:g182 amino-and car 6.1; PID:g182 amino-and car r.; Peat, N.; polymorphic	protein KP39; episia orm; mucin 1 precurso 72-Jun-2000 735; PX0066; S10218;	N 1039 N 1044

QY 545 APTTPKKPAPTTPKEPAPTTPKEPAPT-TTKKPAPTAPKEPAPT 587	462 PPQPIPKAATTTATEVTPQQPIPKAGTDAAPPPAVFKAPSDGRAATPGVPNAATTPQXF 5	OY 492 PTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEP 544	QY 439 TPKEPAPTTKEPAPTTPKEPAPTTAPKKPAPTTPKEPAPTTTKEPS 491 :: :	QY 392 TTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPT 438	Qy 362TKSAPTTPKEPAPTTTKEPAPTTPKEPAPT 391	Qy 311 AETTTKGPALITPKEPTP-TTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTT 361	Qy 253 TVNKETTVETKETTTINKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPK 310	QY 195 VDEAGSGLDNGDEKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSL 252	PKPPV 19 PPP 12	QY 81 AEVHNPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEEH 134	Query Match 9.4%; Score 695; DB 2; Length 1151; Best Local Similarity 25.3%; Pred. No. 9.4e-22; Matches 304; Conservative 108; Mismatches 466; Indels 324; Gaps 59;	Db 1116 SPPPAPVKPPSLP-PPAPVSSPPPVVTPAPPKKEEQSLPPPAES 1158 RESULT 8 RESULT 8 RESULT 8 RESULT 8 RESULT 8 C; Species: Gallus gallus (chicken) C; Species: Gallus gallus (chicken) C; Species: Gallus gallus (chicken) C; Accession: T18535 R; Shimada, K.; Harata, M.; Mizuno, S. J. Cell Sci. 110, 3031-3041, 1997 A; Title: A nuclear matrix-associated high molecular mass nuclear antigen, HMNA, of chick A; Accession: T18535 A; Reference number: Z18955; MUID:9803440 A; Reference number: Z18955; MUID:9803440 A; Accession: T18535 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: mRNA A; Residues: 1-1151 <shi> A; Cross-references: EMBL:D88440; NID:d1177138; PID:d1025045; PIDN:BAA24137.1</shi>	798 1058	0	/ 1
---	---	--	--	--	--	---	---	---	---------------------------	--	---	---	-------------	---	-----

```
RESULT 9
T25697
T25697
Rypothetical protein F16F9.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Accession: T25697
R;Fulton, B.
Submitted to the EMBL Data Library, August 1996
A;Description: The sequence of C. elegans cosmid F16F9.
A;Reference number: Z20071
A;Accession: T25697
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1229 <FUL>
A;Cross-references: EMBL:067956; pIDN:AAB07691.1; GSPDB:GN00028; CESP:F16F9.2
A;Experimental source: strain Bristol N2; clone F16F9
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
A; Gene: CESP:F16F9.2
A; Map position: X
A; Introns: 35/1; 361/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1034
                                                                                                                                                                                                                                                                                                                                                                                                                                 1078 GE 1079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   641
                                                                                                                                                                                                                                                                                                                                                                                               1118 RE 1119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          588 TPK---ETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELA------PTTP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           522 PPTPQSVPSAVTEPKPQPRAAPPPSNEATPAVPSPSPNLKSPLPTIPKPVPLMALTPQPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           873
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | | | :|
VTPRPLPASPVPMAASPASLGPDAARVALATNAASPGAKPEAAGGNGTLMAPMGAANTOM 700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --EEPTPTTP----EEPAPTTP-----KAAAPNTPKE------PAPTTPKEP 664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TAQMVTQLAATKPSPIVPKASPK-ALMTPPPPPPPGLPRALAAKLLGLPSSPVASAMHAK 640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DLRTTPETTTAAPKMT-KETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTL 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AATTP-QPSPM----GAATTLMSPMGAATTPQ---PSPMGAVTTQPPPMAATNTTQPPPM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APTTPKEPAPTTPKGTAPTTLKEP--APTTPKKPAPKELAPTTTKGP----TSTTSDKP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TQMSPMGATTTQMSPMGAAATTQPSPMGAAATQVTATSAGNTMQVSPMGAATPPQTPSVG 820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TLKEPAPTTPKKPAPKELAPTTTKEP-----TSTTSDKPAPTTPKG--TAPTTPKEP 741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APIGAAGAAQTAPMGAAHTHVSPMGAGGATQMSPTGAANTHMSPIGAGGATQMSPMGAAN 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APTTPKEPAPTTPKETAPT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SPOPMSKSPPPDPPKAPSAAAQTSPAAHVANASPGV-TAVSPA---PIGVTEASPSADGA 988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPTTIHKSPDESTPEL--SAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTER 898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APTTPKET----PTTPKEP----APTTPKKP-APTTPETPP-----PTTSEVSTPTTTK 840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVGDGQQQMTPGAAQSVPP-----VTEAAVQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMP 1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RLSPGPTAATDGPKASPAATADVTEAATD--VTAAATAVPA----EAAP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAE 1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -TPKGTAPT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EAAAAAAAAAGAE 1117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---P 1078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -TK 1033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            691
```

```
Query Match
9.1%; Score 676.5; DB.2; Length 1229;
Best Local Similarity 27.2%; Pred. No. 5.7e-21;
Matches 330; Conservative 95; Mismatches 385; Indels 405; Gaps 56;
```

```
밁
                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 328;
      1014 STSPSETA-
                                                               954
                                                                                          920
                                                                                                                     895
                                                                                                                                                 860
                                                                                                                                                                            838
                                                                                                                                                                                                       810
                                                                                                                                                                                                                                    778
                                                                                                                                                                                                                                                               751
                                                                                                                                                                                                                                                                                            719
                                                                                                                                                                                                                                                                                                                     694
                                                                                                                                                                                                                                                                                                                                                  667
                                                                                                                                                                                                                                                                                                                                                                             640
                                                                                                                                                                                                                                                                                                                                                                                                         612
                                                                                                                                                                                                                                                                                                                                                                                                                                     581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                        MNKPEETAKPKDRATNSKATTPKPQKP-----TKAPKKPTSTKKPKTM
                                                      TTSGCSPKTYTTTYPTTTTTSYTTSSTTTTTTYCSTGTNSAGETTSGCSPKTITTTYPC
                                                                               TTEKTTESKITAT --TTQVTSTTTQDTTPFKITTLKTTTLAPKVTT---TKKTITTT-EI
                                                                                                            GT-TVTPSSSKYPGSQTETSVSSTTETTIVPTKTTTSVTTPSTTTTTTVCSTGTNSAGE
                                                                                                                                     PTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETAT
                                                                                                                                                                   PVPTPSSSTTESSSAPVSSSTTESSVAPVPTPSSSSNITSSAPSSI---PFSSTTESFST
                                                                                                                                                                                       PAP----TTPKKPAPTTPETPPPTTSEVSTPT----TTKEPTTTHKSPDESTPELSAE
                                                                                                                                                                                                                          PVPTPSSSTTESSSAPVPTPSSSTTESSVAPVPTPSSSSNITSSAPSSTPFSSSTESSSV
                                                                                                                                                                                                                                                     PT-TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKE
                                                                                                                                                                                                                                                                                 SAPVPTPSSSTTESSSAPVPTPS-SSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSSA
                                                                                                                                                                                                                                                                                                     KEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAP---TTPKEPA
                                                                                                                                                                                                                                                                                                                                                                    PAP----TTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPT-TPKGTAPTTL
                                                                                                                                                                                                                                                                                                                                         PVPTPSSSTTESSSAPVT----SSTTESSSAPVT----SSTTESSSAPVPTPSSSTTESS
                                                                                                                                                                                                                                                                                                                                                                                               PTPSSSTTESSSAPVT----SSTTESSSAP-VPTPSSSTTESSSAPVPTPSSSTTESSSA
                                                                                                                                                                                                                                                                                                                                                                                                                          PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPT-TPEEPTPTTPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                       STTESSSTPVTSS----TTESSSAPVPTPSSSTTESSSAPVPT----PSSSTTESSSAPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTTPKEPSPTTTKEPAPTTPKEPAPT-TPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----PTPSSSTTESSSAPAPTPS---SSTTESSSAPVTSS---TTESSSAPVPTPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STTESSSAPVT----SSTTESSSAPV----PTPSSSTTESSSAPVT---SSTTESSSAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APVTSSTTESSSAPVTSSTTESSSAPVTSS----TTESSSAPVTSSTTESSSAPVPTPSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APTT----TKSAP----TTPKEPAPTTTKEPAPTTPKEPAP----TTTKEPAPTTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---TSSTSESSTTTSSTSESSTTTSSTSESSTSSSTT----APATPTTTSCTKEKPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTP---TTPKEPAPTTKEPAPTTPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APV-----PTPSSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PT--PSSSTTESSS-----APVPTPSSSTT-----ESSSAPVTSSTTESSS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTTTSCTKEKPT--PPHHD------TTPCT----KKKTTTSKTCTKKTTTPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DCDAQCKKYDK----CCPDYESFCAEVHNPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DLSTGCNNYDNQGHSQTDFPGFYWNIDCDNNCGGTKSSTTTSSTSES--STT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
SESTITSPITPVITVVSTTVVTTEYSTSTKPGGEITTTFVTKNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10.9%; Score 809; DB 1; Length 13 27.5%; Pred. No. 2.3e-26; tive 112; Mismatches 526; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 1367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       226;
                           1016
                                                       1013
                                                                                                                                                                                                                           837
                                                                                                                                        919
                                                                                                                                                                     894
                                                                                                                                                                                                                                                                                                           750
                                                                                                                                                                                                                                                                                                                                         718
                                                                                                                                                                                                                                                                                                                                                                     693
                                                                                                                                                                                                                                                                                                                                                                                               666
                                                                                                                                                                                                                                                                                                                                                                                                                            639
                                                                                                                                                                                                                                                                                                                                                                                                                                                       611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              314
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53;
```

망 S В Š В Ş 밁 õ

```
extensin-like protein - maize
C;Species: Zea mays (maize)
C;Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 29-(
C;Accession: S49915
R;Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.
submitted to the EMBL Data Library, June 1994
A;Description: Pex genes: pollen-specific genes with extensin-like (
A;Reference number: S49915
A;Accession: S49915
A;Accession: S49915
A;Accession: S49915
A;Residues: preliminary
A;Molecule type: DNA
A;Residues: 1-1188 < RUB>
                                      Š
                                                                         ДЪ
                                                                                                        Q
                                                                                                                                                                      δÃ
                                                                                                                                                                                                         В
                                                                                                                                          밁
                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                         ₽
                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Cross-references: EMBL:Z34465; NID:g600117; PIDN:CAA84230.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
S49915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 230; Conserv
          944
                                         693
                                                                                                        634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1169
                                                                                                                                          828
                                                                                                                                                                       587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1066
                                                                                                                                                                                                         769
                                                                                                                                                                                                                                          527
                                                                                                                                                                                                                                                                         719
                                                                                                                                                                                                                                                                                                           467
                                                                                                                                                                                                                                                                                                                                          660
                                                                                                                                                                                                                                                                                                                                                                          416
                                                                                                                                                                                                                                                                                                                                                                                                           600
                                                                                                                                                                                                                                                                                                                                                                                                                                            359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             188 PTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSL----PPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1017 PRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           460
                             LKE---PA-----PTTPKK---PAPKELAPTTTKE---PTSTTSDKPAPTTPKGTAPTTP 738
VKSSPPPAPVSSPPATPKSSPPPAPVNLPPPEVKSSPPPTPVSSPPPA---PKSSPPPAP 1000
                                                                                                                               LAPKSSPPHVVVSSPPPVVKSSPPPAPVSSPPLTPKPASPPAHVSSPPEVVKPSTP--PA
                                                                                   TTPKETAP----TTP----KKLTPTTPEKLAPTTPEKPAP----TTPEELAPTTPEEPT
                                                                                                                                                                                            PPAPLSSPPPAPQVKSSPPPVQVSSP-PPAPKSSPPLAPVSSPPQVEKTSPPPAPLSSPP
                                                                                                                                                                                                                              EPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAP
                                                                                                                                                                                                                                                             SP--PKEPVSSPPQTP----KSSPPPAPVSSPPPTPVSSPPALAPVSSPPSVKSSP----
                                                                                                                                                                                                                                                                                       APTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPK
                                                              PTTVISP-PSEPKSSPPPTPVSLPPPIVKSSPPPAMVSSP-PMTPKSSPPPVVVSSPPPT
                                                                                                                                                                                                                                                                                                                                PAKSTPPPEEYPTPPTSVKSSPPPEKSLPPPTLIPSPPPQEKPTPPSTPSKP-PSSPEKP
                                                                                                                                                                                                                                                                                                                                                            PKKPAPTTPKEPAPTT-----PKE---PTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKP
                                                                                                                                                                                                                                                                                                                                                                                               PPAPVASPPPPVKSPPPPTPVASPPPPAPVASSPPPNKSPPPPTPVSSPPPPEKSPPPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPPPPVKSPPPPAPVGSPPPPEKSPPPPAPVASPPPPVKSP--PPPTLVASPPPPVKSPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----LAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPA 358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKV 303
                                                                                                                                                                                                                                                                                                                                                                                                                              PTTTKSAPTTP---KEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTP-KEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APVG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTPHSPPAD------DYVPPTPPVPGKSPPATSPSPQVQPPAASTPPPSLVKLSPPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPIDTVFTRC----NCEGKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TT-----YVTTLAPSAPVTPATN-----AVPTTTTT--TECSAATNAAGET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPKTVTTTVPCSTGTGEYTTEATTLVTTAVTTTVVTTESSTGTNSA-GKTTTGYTTKSVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSDETNICNGKPVDGLTTLRNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTTYLTTIAPTP ---SVTTVTNETPTT-ITTTVCSTGT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          #sequence_revision 12-May-1995 #text_change 29-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----SPPPPVKTTSPPAPIG----SPSPPPPVSVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 784; DB 2;
Pred. No. 2.1e-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----NSAGETTSGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PID: g600118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                domains
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                 885
                                                                                                                                                                                                                                                                                              526
                                                                                                                                                                                                                                                                                                                                                               466
                                                                                                                                                                                              827
                                                                                                                                                                                                                                                                 768
                                                                                                                                                                                                                                                                                                                                                                                               659
                                                                                                                                                                                                                                                                                                                                718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33;
```

Mon

Apr

```
A;Molecule type: DNA
A;Residues: 1-2187 <VOT>
A;Residues: 1-2187 <VOT>
A;Cross references: EMBL:U48363; NID:g1666688; PID:g1666689; PIDN:AAB18732.1
C;Genetics:
C;Genetics:
A;Gene: Naca
A;Map position: 10
A;Map position: 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Introns: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3
A;Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding
C;Keywords: alternative splicing; DNA binding; transcription factor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: T30826
A; Status: prelimina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                    Вb
                                                                                                                                                                                                                                                                                                                                                                                                                δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ф
                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                       밁
밁
                          δõ
                                                            밁
                                                                                       δÃ
                                                                                                                         ф
                                                                                                                                                       Q
                                                                                                                                                                                          밁
                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
Matches 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                       1115
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1058
 1465
                                                                                                                              1348
                                                                                                                                                                                            1294
                                                                                                                                                                                                                                                            1235
                                                                                                                                                                                                                                                                                                                         1175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         145
                                                                                                                                                                                                                                                                                                                                                       507
                                                                                                                                                                                                                                                                                                                                                                                                                       458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    998
                                                                                                                                                                                                                            597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NPTSPPSSKKAPPPSGASQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSS 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QHV---PPTSPPKSPVSDTLSGALTSPPPKGPPATLAETPTYPKKSPKPAASKKTPATPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GDFKYTTPDTSTTQHNK-----VSTSPKITTAKPINPRPSLPPNSDTSKETSL--TVNKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KTPEVTASRLISAVQSPKVDPIMSD-----VTPTSPKKTSATAVPK--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SPT-PPSSKGAPVPSTGA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                  LKETSKSATPGEKSASSPKRSPKTAGPKE---TPPGGVTAVPPEISLPPKETPQNATPNE 1114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PEGVTAVPLEIPPCSKKAPKTAAPKESSATSSSKRAPKTAVSKEIPSKGVTAVPLEISLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKE-PAPTTTKSAPTTPK 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TLATSIPKVTSPSPQKTPKSVSLKGAPAMTSKKAT---EIAASKDVSPSQ--FPKEVPLL
                                                                                                                                                                                                                                                                                                                                                                                     SLAASSQKRSPKTSVPKETPPGGVTAMPLEIPSAPQKAPKTAVPKQIPTPEDAVTILAGS
                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTAPKKPAPTT--PKEPAP----TTPKEPAPTTTKEPSPTTPK----EPAPTTTKSA
                                                                                                                                                                                                             PKKLTPTTPEKLAPTTPEK-PAPTTPEEHAPTTPEEPTPTT-----PEEPAPTTPKA
                                                                                                                                                                                                                                                          APKETPATSSEGVTAVPSEISPSPPTPASKGVPVTLTPKGAPNALAE-SPASPKKVPKTA
                                                                                                                                                                                                                                                                                                                                                          PTTTKEPAPTTT-KSAPTTPK---
PAPTTPKE-PAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTP
                                                                                                                            IPPSPRKAPKTAAPKETPAPS--PEGATTAPVQIPPSPRKGSKKAGSKE-TPTTPSPEGV 1404
                                                                                                                                                          AAPNTPKEPAPTTPKE-PAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTP-----
                                                                                                                                                                                                                                                                                                                         PLSPKKASKTAAPKEAPATPSVGVIAVSGEISPSPKKTSKTAAPKENSATLPPKRSPKTA
                                                              TAAPLEIPISSKKTSKMASPKETLVTPSSKKLSQTVGPKETSLEGATAVPLEIPPSHKKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -TVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -DTSATLSLKSVPAVTSLSPPKAPVAPSNEATIVPTEIPTSLKNALAAATPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TPKEPAPTTPKK-PAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKE 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EPAPTTTKEPAPTTPKEPAPTTTKEPAPTT --TKSAPT-
                                                                                                                                                                                              ---SPQKIPKVAGPKEASATPPSKKTPKTAVPKETSAPSEGVTAVPLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11.0%;
                                                                                               KKPAPKELAPTTTKEPTS----
                                                                                                                                                                                                                                                                                          PAPTTPKEPAP-----TTTKKPAPTAPKEPAPTTPKETAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 812; DB
Pred. No. 2.7e
28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----PPSPKGAPIVPTESSISSKQVPAEILPSPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                          -EPSPTTTKEPAPTTPKEPAPTTPKKPAPTT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2.7e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           449;
                                                                                                 -TTSDKPAPTTPKGTAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 2187;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   257
                                                                                                                                                                                                                                                                                                                                 1234
                                                                                                                                                                                                                                                                                                                                                                                               1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            406
                                                                                                                                                                                                    1347
                                                                                                                                                                                                                                                                   1293
                                                                                                                                                                                                                                                                                                 596
                                                                      1464
                                                                                                                                                                    701
                                                                                                                                                                                                                                    647
                                                                                                                                                                                                                                                                                                                                                                 556
                                                                                                                                                                                                                                                                                                                                                                                                                            506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           acti
```

```
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - yeast (Saccharomyces cerevisiae) N;Alternate names: extracellular glucoamylase; mucin-like protein MUC1; prot. C;Species: Saccharomyces cerevisiae C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Nov-1999 C;Accession: S48478; A26877; B26877; S27281; JC6123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Title: Gene fusion is a possible mechanism A;Reference number: A91831; MUID:87194600 A;Accession: A26877 A;Molecule type: DNA A;Residues: 1-242 <YAM>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
S48478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: EMBL:M16165; NID:g172523; PIDN:AAA35015.1; R;Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, FEBS Lett. 239, 179-184, 1988
A;Title: Similar short elements in the 5' regions of the STA2 & A;Reference number: S27281; MUID:89031230
A;Accession: S27281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Reference number: A; Accession: S48478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Residues: 1-1367 < ROW>
A; Cross-references: GB:Z47047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                submitted to the EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R; Rowley,
                                                                                                                                                                                                                                                        R;Lambrechts, M.G.; Bauer, F.F.; Marmur, J.; proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, A;Title: Mucl, a mucin-like protein that is A;Reference number: JC6123; MUID:96323237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  R;Yamashita, I.; Nakamura, M.; Fuki
J. Bacteriol. 169, 2142-2149, 1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
                                                                                      A;Gene: SGD:MUC1; STA2; MAL5; DEX2; A;Cross-references: MIPS:YIR019c; SG
                                                                                                                                                                          A; Molecule type: DNA
A; Residues: 1-1367 < LAM>
                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: DNA
A; Residues: 1-31 <PAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
A; Residues: 762-1331 <YA2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references:
A; Accession: B26877
                                                                                                                                                                                                                                                                                                                                                 A; Cross-references:
                                                                  A; Map position:
                                                                                                                                                                                                                  A; Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                           A; Accession: JC6123
Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosida; keywords: glycosidase; hydrolase; polysaccharide degradation; transmembrane protein 5-21/Domain: transmembrane #status predicted <TM1>
                                                                                                                                  :Genetics:
                                                                                                                                                  ;Cross-references: GB:U30626; NID:g1304386;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KETAPTTPKEPAPTTPKKPAPTTPETP-----PPTTSEVSTPTTTK-----EPTTIHKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LAAPAV----LPVSSKSPAAPARASASLSPATAAPQTAPKEATTIPSCKKAAATETPIET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTKSPKIPSSKKAPRT--SAPKEFPASPSIK-PVTT--SLAQTAPPSLQKAPSTTIPKEN 1634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDESTPELSAEPTPKALENSPKE-PGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KASGTTASKVPVPAETQEVAVSSRETPVTPAVPPVKNPSSHKKTSKTIELKEAPATLPPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STAPSLEGAPKETSE----TSVSKVLMSSP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTLPSLKEASVLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IBL Data
S48478
                                                                                                                                                                                                                                                                                                                                        EMBL:X13857; NID:g4551; PIDN:CAA32069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL:M16164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Library,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL:238061; NID:g603997; PID:g763364; GSPDB:GN00009; Fukui, S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NID: g172522;
                                                                                             2; SGD:S0001458
SGD:S0001458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      October 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTATSSGKDSHISPVS-DACSTGTTTP
                                                                                                                                                                                                                                                                                                                               J.;
                                                                                                                                                                                                                                                                                         regulated
                                                                                                                                                                                                                                                                                                                Pretorius,
1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    underlying
                                                                                                                                                             PIDN:AAC49609.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -PKKASSSKRASTLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIDN: AAA35014.1;
                                                                                                                                                                                                                                                                                            γd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the
                                                                                                                                                                                                                                                                                            Mss10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         evolution of
                                                                                                                                                                                                                                                                                                                                                        PID:94552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and
                                                                                                                                                               PID: g1304387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PID:g172526
A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PID: g172525
                                                                                                                                                                                                                                                                                              is critical for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       from
                                                                                                                                                                                                                                                                                                pseud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sacc
```

C; Keywords: glycosidase; hydrolase F;5-21/Domain: transmembrane #stat F;1350-1366/Domain: transmembrane

#status

predicted <TM2>

```
В
                                        Ş
                                                                           밁
                                                                                                       Qy
                                                                                                                                          밁
                                                                                                                                                                        Ş
                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                       ρy
                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       . Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 맑
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
1207 -PKPKEVSKEPPKPTEPPKP-AAPKKWKPPWEDDPDEPEADFTMPAPKKPDTEDP 1259
                                                                         1171
                                                                                                                                            1111
                                                                                                                                                                                                             1051
                               994 TEKPOKPTKAPKKETSTKKPKIMPRVRKPK-----TIPTPRKMTSTMP 1036
                                                                                                                                                                                                                                            843
                                                                                                                                                                                                                                                                                                               794
                                                                                                                                                                                                                                                                                                                                               938
                                                                                                                                                                                                                                                                                                                                                                                752
                                                                                                                                                                                                                                                                                                                                                                                                                883
                                                                                                                                                                                                                                                                                                                                                                                                                                               693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          654 AEPNSPVVP-PTPVKNPVKKWKPPWEDDDAPAKPVSLPEPEKK-TPVLAKKAPTKPDSEA 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   306
                                                                  PSKKPDTEDPADPLG------GPKTKDPK-----LNKKAPAEKPTEK-----
                                                                                                                                 LSKKAPVEKPKPTTDPKDDKLKPSPAKKPEKAPEPAAPKKWKPVWDDDPDEPEADFTVPA 1170
                                                                                                                                                                   AT-----KPEMTTTAKOKTTERDLRTTPETT--TAAPKMTKETATTTEKTTESKITATT 933
                                                                                                                                                                                                    VKKWKPPWEDDDEPSEPVSAPEPEKKTPVLAKKAPTKPATKPDSEAAADPVSGPTSKDPK 1110
                                                                                                                                                                                                                                                                                                                                      PSPKKAVPE--KEPAKVAAKPRDLSPKKAIPIPANTQEAPPTPVKNPVKKWKPPWEDDDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAPTTPKETAPTTPKKLTPTTP-----EKLAPTTPEKPAPTTPEELAPTTPEEP-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AADPVSGPSSKDPKLAKKAPVKPRDPSPMKAVPIKPAPKT--EVPPAVVKKPEPVAKSRD
                                                                                                                                                                                                                                    TTIHKSP----DESTPELSA-EP---TPKALENSPKEPGV-------PTTKTPA 881
                                                                                                                                                                                                                                                                       PAEPVSAPEPEKKTPVLAKKAPAKPRDP---SPKKAAPVAAK-PDPKIPEV-PPTPVKNP
                                                                                                                                                                                                                                                                                             P----APTTPKET-----APTTPKEPAPTTPKKPAPTTPETPPTTTSEVSTPTTTKEP 842
                                                                                                                                                                                                                                                                                                                                                                     TTPKGTAPTTLKEPAPT-----TPKKPAP-----KELAPTTTKGPTS-----TTSDK 793
                                                                                                                                                                                                                                                                                                                                                                                                       VKNPVKKW-KPPWEDDDEPTEEVKKPSE--PEKKTPVLAK-KEPEKPKD-APKVAAKPRD
                                                                                                                                                                                                                                                                                                                                                                                                                                      LKEPAPTTPKKPAPKELAPT-TTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PVKPRDPSPKKAVPAKPSTKTDAPPVSVKKPEPVSKPKEPSPKKAEPNSPVVP----PTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTT-PKEPAPTTPKETAPTTPKGTAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DDVSKPTDAKKTPSLAKKDPAPAKESLKPKADTKAPAKPRDPSPKKVAPTAPEKKTPVLA 593
                                                                                                    TQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKAT 993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PSPKKAK-AEPNSP--VVPPTPVKNPVKKWKPPWEDDDAPAEPVNVPEPEKKTPVLAKKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPTTTKEPAPTTPK--EPAPTTPKKPAP--TTPKEPAPTTPKEPAPTTTKKPAPTA-PKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPTTPKEPAPTTKEP----SPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTP-KEP 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KKEPAGPADSKTKEPEKSKPRDPSPKKAVPAKPVPKTEVAPAAVKKPEPISKPKDTAPKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KE----PAPTTPKEPTPTTPKEPAPTTKEPAPTTPK-EPAPTAPKKPAP-----TTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----PTTEKEPAPTTKEPAPT-----TTKSAPTTPKEP-----APTTPKKPAPTTP 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPATSKTPETPEKKDPVKPRDSSPKKVAAKPDSAQAPA-TPVKNPVKKWRPPWEDDETPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QADPKAQEVPPTPVKNPVKKYKPPWEVDDEDPVEEVKQPEAPAKKTPVLKRKEPAAKDTA 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EP----TP-TTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPA---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KPTPKAE----TTTKGPA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              340
                                                                      1206
                                                                                                                                                                                                                                                                                                                                        995
                                                                                                                                                                                                                                                                                                                                                                                                         937
                                                                                                                                                                                                                                                                                                                                                                                                                                        751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             533
```

nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - m R;Alternate names: alpha NAC protein C;Species: Mus musculus (house mouse) C;Pate: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 24-Nov-1999 C;Accession: T30826 R;Yotov, W.V.; St-Arnaud, R. Genes Dev. 10, 1763-1772, 1996 A;Title: Differential splicing-in of a proline-rich exon converts alphaNAC into A;Reference number: Z20889; MUID:98312450 mouse ø

mus

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         먑
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DЪ
망
                                                                  δõ
                                                                                                                                                          망
                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                       Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       뮍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 312; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1397
        1601
                                                                                   1157
                                                                                                                                                                                                                                                                                                                                 1499
                                                                                                                                                                                                                                                                                                                                                                                                            1084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TPSDEPTPSE----TPEEPIPTDTPSDEPTPSDEPTPSDEPTPSD-EPTPSDEPTPSETP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SDEPTPSDEPTPS--DEPTPS-TPSETPSETPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KPAPTTPKEPAPTTPKEPTTPKEPAPTTKEPAPT-TPKEPAPTAPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TKSAPTTPKEPAPTTTKEPAPTTPKEPA----PTTTKEPAPTTTKSAPTTPKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VVIQPAPIKAASDEPIPTDTPSDEPTPS----DEPTPS----DEPTPSDEPTPSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EEPAPTTPKAAAPNTPKEPAPTTPKEPA--PTTPKEPAPT-TPKETAPTTPKGTAPTTLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTPKEPA--PTTTKEPSPT-TPKEPAPTTTKS-APTTTKEPAPT---TTKSAPTTPKEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSD----EPTPSDEPTPSDEPTPSDEPTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPS--DEPTPSETPEEPIPTDTPSDEPTPSDEPT----PSDEPTP--SDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAPTTPKKPAPT-TPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKGTAPTTLKEPAPT-TPKKPAPKELAPTTTKGPTSTTSDKPAPT---TPKETAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APT---TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPT---TPEELAPTTPEEPTPT-TP
        YIKRSESV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YLRGYPDGSFRPERNITRAEAAVIF-----AKLL----GADESYGAQSASPYSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGGSGGGGGGGGGGTVPTSPTPTPTSKPTSTPAP---TEIEEPTPSDVPGAIGGEHRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATNSKAT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEPTPSDEPTPSETPEEPTP----TTTPTPTPSTT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKEPGVPTTKTPAAT - - KPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---DEPTPSDEPTPSETPEEPI-----PTDTPSDEPTPSDEPTPSD-EPTPSDE
                                                                                   PARRITEVWGIPSPIDTVFTRCNCEGKTFFFKD--SQYWRFTNDIKD
                                                                                                                                                                  TIDISNPKFDD----CVGHWAQEFIEKLTSLGYISGYPDGT--
                                                                                                                                                                                                                ---IINPMLSDETNICNG----KPVDGLTTL-----RNGTLVAFRGHYFWMLSPFSPPS
                                                                                                                                                                                                                                                                                                                            LADTHWAAWAIKFATSQGLFKGYPDGTFKPDQNITRAEFATVVLHFLTKVKGQEIMSKLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12.8%; Score 950; DB 2; ilarity 31.0%; Pred. No. 4.5e-32; Conservative 104; Mismatches 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHM-
        -ALINRALERGPLNGAPKLEPDVNESYWAF-GDIMD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --TPKPQKPTKAP-KKPTSTKKPKTMPRVRKPKTTPTPRKM------
                                                                                                                                                                                                                                                                                                                                                                                                                ----VFMPE-----VTPDMDYLPRVPNQGI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                361;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 230;
                    1642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----PTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      979
                                                                                                                                                                                                                                                              1156
                                                                                                                                                                                                                                                                                                                                         1558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
```

```
A;Reference number: Z20986
A;Accession: T31108
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1489 <GOE>
A;Cross references: EMBL:AF061185; NID:g3851513; PID:g3851514; C;Genetics:
A;Gene: car90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cyst germination specific acidic repeat protein precursor - Phytophthora infestans (;Species: Phytophthora infestans (potato late blight agent) C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999 (;Accession: T31108 R;Goernhardt, B. submitted to the EMBL Data Library, April 1998 A;Reference number: Z20986
                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QY
                                                                                                                                                                                                                                                             Db
                                                                                                                                                                                                                                                                                             ΔÃ
                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                             Qy
                                                                                                                                                                                    Вþ
                                                                                                                                                                                                                        Qγ
δõ
                                     DЬ
                                                                    Qy
                                                                                                              Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
Matches 368; Conserv
                                                                                                                                                                                                                                                                                                                                                                               477
   713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     YAPTEETTYAPTEETPYEPTEETTYAPTKETTYAPT---EETTYASTEETTYAPTEETTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTT-----QHNKVSTSPKITTAKPI 234
                                                                                                                                                                                                                          T--TPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTT
                                                                                                                                                                                                                                                                                               --APT-TTKSAPT--TPKEPSPTTTKEP-----APTTPKEPAPTTPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                         TTPKEPAPT--TTKEPSPTTPKEP-----APT-TTKSAPT-----TTKEP----
                                                                                                                                                                                                                                                                                                                                                                                                                 TYAPTEETTYASTEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETTYAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EETTYAPTEETT-YAPTEETMYAPIEETTYAPTEETTYAPAEETPYEPTEETTYAPTEET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTTTKEPAPTTTKSAPT-----TPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APAEETPYEPTEET-TYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPAEETPYEPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EPASTTPKEPTPTTIKSAPTTPKEPAPT-TTKSAPTTPKEPAPT--TTKEPAPTTPKEPA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EETTYAPTEETTYAPTEETTYAPTEETPYEPTEETTYTPTEETTYAPTEETTYAPTEKTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REVTPYAPTEKPYDVEETTYVTEESTYAPTKSETNAPTERMHYAHIEKPCDTEVTMYAPT
                                                                                                                                                                                                                                                             TYAPTEATTYAPTEETPYAPTEETTYEPTGETTYAPTEETTYAPTEETTYAPTEETTYAP
                                                                                                                                                                                                                                                                                                                                          TEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETMYAPIEETTYGPTEET
                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKEP----APTAPKKPAPTTPKEPAPTTPKEP----APTAPKKPAPTTPKEPAP
                                                                                                                                                                                          TEETPYEPTEETTYAPTEETPYEPTEET--TYTPTEETTYAPTEETTYAPTEKTTYAPTE
                                                                                                                  ETTYAPTEETPYEPTEETTYAPTKETTYAPTEETTYASTEETTYAPTEETTYAPAEETPY
                                                                                                                                                   PEELAPT -- TPEEPTPTTPEEP-----APT-----
 -TTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEP---
                                     EPTEETTYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPAEETPYEPTEETTYAPTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                        -APTTPKEPAPTTPKETAPTTPKGTAPT---TLKEPAPTTPKKPAPK-ELAPT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------PNSDT----SKETSLTVNKETTVETKETTT---TNKQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11.7%; Score
31.7%; Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; score 863.5; DB 2;
; pred. No. 1.4e-28;
55; Mismatches 509;
                                                                                                                                                         TPKAAAPNTPKEPAPTTPK
     APTTPKGTAPT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIDN: AAC72308.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      394
                                                                                                                                                                                                                                                                                                                                                                                       513
                                                                                                                                                                                                                                                                                                                                                                                                                                                              476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          570
                                                                                                                                                                                                                                                                                                            562
                                               987
                                                                                                                       927
                                                                                                                                                                                              867
                                                                                                                                                                                                                                    620
                                                                                                                                                                                                                                                                           809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        51;
```

```
J. Clin. Invest. 87, 77-82, 19
A. Title: Human bronchus and il
A; Reference number: A61257; M
A; Accession: A61257
A; Status: not compared with co
A; Molecule type: mRNA
A; Residues: 'T', 1925-1948, 'Tri
A; Experimental source: bronch
A; Xu, G.; Huan, L.; Khatri, I.
                B
                                               Š
                                                                                          밁
                                                                                                                         Š
                                                                                                                                                                    В
                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                       망
                                                                                                                                                                                                                                                                                                                                                                                                                                            ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A/Cross-references: GDB:120203; OMIM:158370
A:Map position: 11p15.5-11p15.5
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology;
C:Keywords: glycoprotein; intestine; tandem repeat
F;2766-2834/Domain: von Willebrand factor type C repeat homology <vWC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           R:XU, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Forstner Blochem. Blophys. Res. Commun. 183, 821-828, 1992
A:Title: Human intestinal mucin-like protein (MLP) is homologous with rat MiP in the C-R:Reference number: P00328; MUID:92198477
A:Reference number: P00328; MUID:92198477
A:Rocession: P00328
A:Residues: 2328-2468 <XUG>A:Residues: 2328-2468 <XUG>A:Experimental source: small intestine
A:Residues: P00329
A:Recession: P00329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Gene:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: protein
A: Residues: 2328-2342, 'K', 2344-2354 < XUG1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local s
Matches 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Molecule type: mRNA
Residues: 'T',1925-1948,'TTS',1952-1954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Status: not compared with conceptual translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Experimental source: bronchus
                1561
                                                                                          1502
                                                                                                                                                                    1444
                                                                                                                                                                                                                                                  1384
                                                                                                                                                                                                                                                                                                                           1324
                                                                                                                                                                                                                                                                                                                                                                                                       1272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1079
                                                   463
                                                                                                                               404
                                                                                                                                                                                                                                                                                       297
                                                                                                                                                                                                                                                                                                                                                                  279
                                                                                                                                                                                                                                                                                                                                                                                                                                              236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 27.4 tes 378; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GDB:MUC2
TPSPPPTTTTPPPTTTPSPPTTTTPSPPTITTTPPPTTTPSPPTTTTPPPTTTPPPTTTPSP 1620
                      PKKPAPTTPKEPAPTTTKEPSPTTPKEDAPTTTKSAP-TTTKEDAPTTTKSA 521
                                                                           PPMTTPITPPASTTTLPPTTTPSPPTTTTTPPPTTTPSPPTTTPITP-PTSTTTLPPTT 1560
                                                                                                              AP-TTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTA 462
                                                                                                                                                       TTTTPLPTTTPSPPISTTTPPPTTTPSP-PTTTPSPPTTTPSPPT-TTTTPPPTTTPS
                                                                                                                                                                                                                                 IRVNCCWPMDKCITTPSPPTTTTPSPPPTTTTTLPPTTTTPSPPTTTTTPPPTTTPSPPIT 1443
                                                                                                                                                                                          TTIKSAP-TTPKEPAPTTIKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKS 403
                                                                                                                                                                                                                                                                                                               EPFDGVCGAPEDIECRSVKDPHLSLEQHGQKVQCDVSVGFICKNEDQFGNGPFGLCYDYK
                                                                                                                                                                                                                                                                                                                                                                                                                                PRPSLPPNSDTSKETSLTVNKETTVETKET--TTTNK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CGCYVEDTHYPPGASVPTEETCKSCVCTNSSQVVCRPEEGK-----
                                                                                                                                                                                                                                                                                                                                                            KTTS--AKETQSIEKTSAKD-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NRTKKKPTPKPPVVDEAGSGLDNGDF---KVTTPDTSTTQH---NKVSTSPKITTAKPIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E--CEWHYEPCGNRSFETCRTINGIHSNISVSYLEGCYPRCPKDRPIYEEDLKKCVTADK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ESFCAEVHNPTSPPSSKKAPPPSGASQTI-----KSTTKRSPK--PPNKKKTKKVIESEE 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SCSCDTGGDCECFCSAVASYAQECTKEGAC-----VFWRTPDL-CPIFCDYYN-+-PPH 1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TCNCDY -- NCQHYMECCPDFKRVCTAELSCKGRCFESFERGRECDCDAQCKKYDKCCPDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NOTO----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----ITEEH----SVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKK 181
                                                                                                                                                                                                                                                            -----LAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTP
                                                                                                                                                                                                                                                                                                                                                                                         -TSFTTTTTTTTPTSSTVLSTTPKLCCLWSDWINEDHPSSGSDDGDR 1323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      intestine express the MUID: 91086481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     107; Mismatches 514;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -DGAFCYWEICGPNGTVEKHFNICSITTRPSTLTTFTTI 1271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1090; DB 2;
Pred. No. 1.4e-37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <JAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         same
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3020;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 382;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                  -QTSTDGKE 278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                       1501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1186
                                                                                                                                                                                                                                                                                                                                                      296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yon
```

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1664 <FUJ>
                                                                                                                                                                       S-layer protein - Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18262
            A; Cross-references:
                                                                                A; Reference number: Z18847; A; Accession: T18262
                                                                                                                                                 R;Fujino, T.; Beguin, P.;
J. Bacteriol. 175, 1891-18
                                                                                                                                 A; Title: Organization
                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       망
                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δõ
                                                                                                                                                                                                                                                                                                                                                                                                2311
                                                                                                                                                                                                                                                                                                                                                                                                                                 1075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2206
                                                                                                                                                                                                                                                                                                                             2364 D
                                                                                                                                                                                                                                                                                                                                                                1112 N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1852 NEDQKPGGVIPMAFCLNYEINVQCCECVTQPTTM---TTTTTENPTPTPITTTTTVTPTP 1908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1792 LDSGKPNFHKPGGDTELIGDVCGPGWAANISCRATMYPDVPIGQLGQTVVCDVSVGLICK 1851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              522
                                                                                                                                                                                                                                                                                                                                                                                     GCCWHWECDCYCTGWGD-PHYVTFDGLYYSYQGNCTYVLVEEISPSVD-----NFGVYI
                                                                                                                                                                                                                                                                                                                                                                                                                    G------AEGETPHMLL------RPHVFMPEVTPDMDYLPRVPNQGIII 1111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SAWTPTPTPLSTPSIIRTTGLRPYPSSVLICCVLNDTYYAPGEEVYNGTYGDTCYFVNCS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MTSTAPPSTPTAPTTTSGGHTLSPPPSTTTSPPGTPTRGTTT--GSSSAPTPSTVQTTTT 2145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TSTTSDKP----APT-TPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKK 772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNP-----KSEDAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LSCTLEFYNWSCPSTPSPTPTPSK-STPTPSKPSSTPSKPTPGTKPPECPDFDPPR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTQDTTPFKITT----LKTTTLAPKVTTTKKTITTTEIMNKPEE------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAATKPEMTTTAKOKTTEROLRITPETTTAAPKMTKETATTTEKTTESKITATTTQVTST 939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TPTPTTGSTGPPTHTSTAPIAELTTSNPPPESSTPQTSRSTSSPLTESTTLLSTLPPAIE 2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPPPTTSEVSTPTTTK-----EPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKT 879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTTVLITTTTMTPTPTPTSTKSTTVTPITTTTTVTATPTPTGTQTPMIPISTTTTVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAPKELAPTTTKGPTST-TSDKPAPTTPKETAPTTPKEPAPTTPKKP-----APTTPE 824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TPTSTQSPNGLQAPTPTPISTTTTVTPTPTPTGTQTPT-TTPITTTTTTVTPTPTPTGTQT 1967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTP----KETAPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTTTPITPTSTTTLPPTTTTPSPPTTTTPPPTTTPSPPTTTTPSPPITTTTPHPPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SPTTTPSPPTTTMTTLPPTTTSSPLTTTPLPPSITPPTFSPFSTTTPTTPCVPLCNWTGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PSSPITTTP--SPPTT----TMTTPS--PTTTPSSPITTTTPSSTTTP-SPPPTTMTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P-TTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTTKKPAPTA
                                                                                                                                                                                                                                                                                                                             2364
                                                                                                                                                                                                                                                                                                                                                            1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                       -QENETWWLCDCFM-ATCKYNNTVEIVKVECEPPPMPTCSNGLQPVRVEDPD
EMBL:X67506; NID:g296879; PID:g296881; PIDN:CAA47841.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ----TAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                                             1899, 1993
a Clostridium thermocellum
                                                                                                                                                Aubert, J.P.
899, 1993
                                                                                              MUID: 93209931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----TAPTTLKEPAPTTPKKPAPKELAPTT--TKEP
                                                                                                                          gene
                                                                                                                      cluster encoding the cellulo
                                                                                                                                                                                                                                                                                                                                                                                                                                                         2310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1074
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                685
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  640
```

```
Run
                                                                                                                                                                                                                                                                                                                                                                                              Minimum
Maximum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OM protein
                                                                                                                                                                                                                                          Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence:
pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                               BB BB
                  559
559
                                                             695
676.5
666.5
663.5
643.5
633.6
632.5
627.5
607.5
556.5
551
551
                                                                                                                                                                                                                                           Score
                                      605
591.5
568.5

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                              seq
                                                                                                                                                                                                                                                                                                                                                                                              length:
length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-556-246-1_COPY_25_1404
7410
                                                                                                                                                                                                                                            Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            April 26, 2002, 16:18:10;
                                                                                                                                                                                                                                                                                                                                                                                                                                     219241 seqs, 76174552 residues
                                                                                                                                                                                                                                                                                                                                               PIR_68:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GenCore version Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                                                                                                                Listing
                                                                                                                                                                                                                                                                                                                                                                         Maximum
  QDLSSCAGRCGEGYSRDATC.....ARAITTRSGQTLSKVWYNCP 1380
                                                                                                                                                                                                                                                                                                                 pir3:*
pir4:*
                                                                                                                                                                                                                                                                                                                                 pir2:*
                                                                                                                                                                                                                                                                                                                                        pir1:*
                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                2000000000
 1664
1489
1274
2187
1367
                                                                                                                                                                                                                                                                                                                                                                 first
                                                                                                                                                                                                                                                                                                                                                                         Match
                                                                                                                                                                             1188
                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                 100%
45 summaries
                                                                               T29757
T19431
F75518
                                        S50125
T34434
A53577
                                                                                                                                        T45025
                                                                                                                                                                      T18535
                                                                                                                                                                               S49915
                                                                                                                                                                                              T30826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search time 82.85 Seconds (without alignments) 1268.810 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Compugen Ltd
                                                                                                                hypothetical prote
nucleolar phosphop
extensin class 1 p
                                                                                                                                               mucin i precursor, elastic titin - hu
                                                                                                                                                                       extensin-like prot
high molecular mas
                                                                                                                                                                                        glucan 1,4-alpha-g
                                                                                                                                                                                             nascent polypeptid
                                                                                                                                                                                                                       s-layer
                                                                                                                                                                                                                                              Description
                                                                 hypothetical prote
gene LF3 protein
                                                                                                 hypothetical prote
protein UNC-89 - C
                                                                                                                                        mucin MUC5B,
                                                                                                                                                               hypothetical prote
                                                                                                                                                                                                       hypothetical prote
                                                                                  hypothetical
                   hypothetical
                                  hydroxyproline-ric
                                                 hypothetical
                                                                                         hypothetical
                                                          larval glue protei
                                                                                                                                                                                                                germination s
                                                                                                                                                                                                                         2 precursor,
er protein -
                                                                                                                                                                                                                        protein
                                                                                         prote
                                                                                                                                          trach
                                                  prote
                                                                                  prote
```

ALIGNMENTS

A;Cross-references: GB:M94132; NID:g186397; PIDN:AAA59164.1; PID:g186398
A;Experimental source: colon
A;Note: sequence extracted from NCBI backbone (NCBIP:116698)
R;Toribara, N.W.; Gum Jr.; J.R.; Culhane, P.J.; Lagace, R.E.; Hicks, J.W.; Petersen, J. Clin. Invest. 88, 1005-1013, 1991
A;Title: MUC-2 human small intestinal mucin gene structure. Repeated arrays and polym A;Reference number: A43932; MUID:91358717
A;Accession: A43932 R;Gum Jr., J.... Biol. Chem. A; Molecule type: DNA
A; Note: sequence inconsistent with the nucleotide translation
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence extracted from NCBI backbone (NCBIN:55749, NCBIP:55750)
A; Note: sequence inconsistent with the nucleotide translation
A; Note: sequence extracted from NCBI backbone (NCBIN:AAA59875.)
A; Note: sequence extracted from NCBI backbone (NCBIN:AAA59 A; Molecule type: mRNA A; Residues: 2037-3020 <GU3> A;Note: sequence extracted from A;Accession: B45106 A; Cross-references: GB:M94131; NID:g186395; PIDN:AAA59163.1; PID:g186396 A; Note: sequence extracted from NCBI backbone (NCBIP:116706) A; Molecule type: mRNA A; Residues: 626-1895 <GU2> A; Status: not compared with conceptual translation A; Reference number: A45106; A; Accession: A45106 A; Title: The human MUC2 intestinal mucin has A; Reference number: A45106; MUID:93016075 A; Cross-references: GB:L21998 A; Molecule type: mRNA A; Residues: 1-639 <GU1> A; Accession: A49963 A; Status: not compared with conceptual translation A;Cross-references: GB:M22405; NID:g188873; PIDN:AAA36334.1; PID:g188874 A;Experimental source: intestine A; Molecule type: mRNA A; Residues: 1916-2193 Accession: B33532 J.R.; Gallup, 267, Hicks, mRNA 21375-21383, 1992 M.W.; Yan, P.S.; Gum, <GU4> Toribara, N.W.; Rothe, E.M.; Lagace, R.E.; Kim, J.R.; cysteine-rich subdomains located both up Kim, Y.S.; Sequence analysis and Basbaum, C.B <TOR> evi

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
US-08-276-967-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Дb
Ş
                                     д
                                                                    S
S
                                                                                                             ₽
                                                                                                                                            δÃ
                                                                                                                                                                                 В
                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: US-08-276-967-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 2, Application US/08276967 Patent No. 5851817
                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 163; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: USA
COUNTRY: USA
ZD19: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, VG
COURRENT APPLICATION DATA:
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: 713-789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENCTH: 2476 amino aci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UT:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713,787-1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Garbers, David L.
TITLE OF INVENTION: Species-Specific Egg-Binding Proteins
TITLE OF INVENTION: Sperm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION: NAME: Kitchell, Barbara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    869 -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPXPGSTAPXAHGVTSAPDX 924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 824 KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT-SRIAEAMLQTTTRPNQ 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              821 APDXRPXPG-----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                925 RP 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 883
                                                                                                                                                                                                                        406 PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEP 465
581 GTAPTTLKEPAPTTPKKPAPKELAPT-TTKGPTSTT---SDKPAPTTPKETAPTTPKEPA 636
                                        419
                                                                                                                                                466 APTTPKAAAPNTPKEPA----PTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEP 521
                                                                                                                                                                                       313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: US/08/276,967 FILING DATE: Submitted Herewith CLASSIFICATION: 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: P. O. CITY: Houston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE: /13-789-2679
                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ďΤ
                                                              APTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPT-TPKEPAPTTPKEPAPTTPK 580
                                                                                                               TIPTEKSTVPT--KKPTVFKEPTLPPE-GPTVPAE-RPTTPPEGPAVPPKG--PTVLTE- 418
                                                                                                                                                                                   PSETSVSTEKPVAPTE----KPTVPSEIYTIPTEKPMVHMEKPIVHT--EKPTVPT-EKP 365
                                     -----WPTSHTEKSTVHTEKPILPTGKSTIPTEKPMVPTKRT-----TTP- 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Texas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               E: Arnold, White & Durkee P. O. Box 4433
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2476 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hardy, Daniel M
                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                               single
                                                                                                                                                                                                                                                                                  8.4%; Score 419.5; DB 2; 32.5%; Pred. No. 8.5e-19;
                                                                                                                                                                                                                                                                 43; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   UTSD:418\KIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Version #1.25
                                                                                                                                                                                                                                                                                                      Length 2476;
                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                   119;
                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      868
                                                                                                                                                                                                                                                                       31;
```

```
Qy
                                                                                                                                                                                                           Вþ
                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                      Qy
Db
                               Qy
                                                                 Вb
                                                                                                 УŽ
                                                                                                                                       밁
                                                                                                                                         614
                                                                                                                                                                          747
                                                                                                                                                                                                             572
                                                                                                                                                                                                                                                                                                                  637
                                                                                                                                                                                                                                                                                                                                                    459
 694 AHFERCACPVSCQ-SPTPNCEL 714
                                                                                                      807
                                                                                                                                                                                                                                                                             VPTEKPIVPTEKHTIPTEKLTVLTERTTTPTERTTIPTEKPTVPTEKPSVPT-EKPTVPT
                                                                                                  KPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT 866
                                                                                                                                       ERTT-----TPTIRTTTPTERTT---IPTKKTT-----VPTEKTIIPT-----ERTI 652
                                                                                                                                                                      EKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETA 806
                                                                                                                                                                                                                                                                                                                PTTPKKPAPTTPETPPP-----TTSEVSTPT-----TTKEPTTIHKSPDESTPELSAEPT 686
                                SRIAEAMLQTTTRPNQTPNSKL 888
                                                                                                                                                                                                                                          PKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTT 746
                                                                                                                                                                                                                                                                                                                                                   -TERTTIPAEKPTV---PIEKPMVPTERTTIPTERTTIPTEKPTVPTEKLTVPT--EKPI 512
                                                                                                                                                                                                             ----EEPTIPTEKLTVPT--ERTTTPTKRTTTPTIRTTTPTI-----
                                                                 -TTPQP-SPTLVPTQPAAVVMPST-----SATTVTPRTTIASCP---PN 693
                                                                                                                                                                                                                  ---RTTTPT
                                                                                                                                                                                                                    613
                                                                                                                                                                                                                                                                                        571
```

Search completed: April 26, 2002, 16:16:35 Job time: 186 sec

```
PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: US 08/039,320

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/403,576

PRIOR APPLICATION NUMBER: US 08/403,576
                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: (703) 836-202
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL
                   FEATURE:
                                                                                                                                                                                                                       LOCATION: 128..1899
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Eloppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
                                 LOCATION: 144
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                   FEATURE:
                                                                                                                                                                                                                   FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE/DOCKET NUMBER: 01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SOFTWARE: Patentin Rel
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
                                                                                                                             OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                               TYPE: amino acid
STRANDEDNESS: sin
TOPOLOGY: linear
   NAME/KEY:
                                                                                                NAME/KEY:
                                                                                                                                                                                                                                                                                                                  NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: WO POFILING DATE: 23-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: United States ZIP: 22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: Alexandria
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2, Application US/09083116
o. 6203795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                           2035 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P.O. Box 1404
                                                                                                                                                                                                                                                                                              Peptide
128..1899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HAREUVENI, Mara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KIENY, Marie-Paule
Peptide
                                                                                                Peptide
                                                                                                                                                                                                Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LATHE, Richard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAMBON, Pierre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BURNS,
                                                                                                                                                                                                                                                                                                                                               peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                            3) 65.
) 836-2021
NO: 2:
                                                                                                                                                                                                                                                                                                                                                                             single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PHARMACEUTICAL COMPOSITION FOR THE TREATMENT OR PREVENTION OF A MALIGNANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOANE,
                          /note= "Amino acid 144 is Y = Xaa which is the codon for Thr or Asn or ACG; and Asn = AAT or AAC."
                                                                                                                        /note= "Amino acid 134 is X1 = Xaa
Xaa Xaa which is the codon for Pro or
CCC, CCA, or CCG; and Ala = GCT, GCC,
                                                                                                                                                                                                                       /note= "The amino acids spanning 128 to 1899 constitute a repeated region wherein the 20 amino acids, 17 of which are fixed. The number of repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO PCT/FR91/00835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US/09/083,116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08/479,537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              017753-025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SWECKER & MATHIS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Version
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 L.L.P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                . 30
                                         wherein Thr -
                                                                                                                      Ala wherein GCA, or GCG
                                        ACT,
                                                                                                                                 Pro -
                                        ACC,
                                                                                                                                                                                                                                        repeat
                                                                                                                                        CCT,
                                        AC
```

```
Ş
                                      Вb
                                                              Ωy
                                                                                              망
                                                                                                                         Q
                                                                                                                                                         Ъ
                                                                                                                                                                                  δÃ
                                                                                                                                                                                                                    Ъ
                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                          Вþ
                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          뭥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-083-116-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 24.4
Matches 235; Conservative
       764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURE:
                                                                                               703
                                                                                                                                                                                                                    586
                                                                                                                                                                                                                                         572 KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                      531 TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS----APDXR
                                                                                                                                                                                                                                                                                                            524
                                                                                                                                                                                                                                                                                                                                          472
                                                                                                                                                                                                                                                                                                                                                                       469
                                                                                                                                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                                                                                                                                                  413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCATION: 147
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        134 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION: 1..21
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16 TVLTV------VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75
                            DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                       DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP
                                                                                                                                                                                                            PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX
TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ
                                                         -TKPEMTTTAKOK----TTEROLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                 RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                  KPAP -- TTPK ---- ETAPTTPKEPAPTTP ---
                                                                                                                                                                                                                                                                                                    TTP------KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP 571
                                                                                                                                                                                                                                                                                                                                 APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                             TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP 523
                                                                                                                                                                                                                                                                                                                                                                                            SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                                                                          TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKE--PAPTTKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTPKEPAPTTT-----KSAPTTPKEP-----APTTTKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
                                                                                                                      ---IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note- "?
21 amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note= "Amino acid 147 is X2 = Xaa which is the codon for Pro or Ala wherein or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Amino
no acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 3.9e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             precursor
                                                                                                                                                                                 ---KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     461;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              o 21 are a
sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pro •
                                                         763
                                                                                                                                                                                                              644
                                                                                         762
                                                                                                                      707
                                                                                                                                                 702
                                                                                                                                                                               661
                                                                                                                                                                                                                                          617
                                                                                                                                                                                                                                                                       585
                                                                                                                                                                                                                                                                                                                                  530
                                                                                                                                                                                                                                                                                                                                                                                                                            468
                                                                                                                                                                                                                                                                                                                                                                                                                                                         419
                                                                                                                                                                                                                                                                                                                                                                                             471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CCT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      င္ပင္ပင
```

235;

Conservative

63;

Mismatches

461;

Indels

203;

Gaps

48;

```
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTUNAL ARTHUMALER: TESKALL,
NAME: TESKALL,
REGISTRATION NUMBER: 01//-
REFERENCE/DOCKET NUMBER: 01//-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-620
TELEPHONE: (703) 836-2021
TO NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: (703) 836-20: INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 90
FILING DATE: 23-OCT-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: Patentin Rel
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE: 04-APR-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                          LOCATION: 128..1899
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                 FEATURE
                                                           NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                             LOCATION: 147
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STATE: V
                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: amino acid
STRANDEDNESS: sir
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION NUMBER: FILING DATE: 07-JUL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM:
                                                                                                                                                                                                                                            NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: WO PCT/FR91/00835 FILING DATE: 23-OCT-1991
                                                                                                                                                                                                                                                                                 OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                        OTHER INFORMATION: OTHER INFORMATION:
                                                                                                                                                                                                                                  OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                      LOCATION:
                                                                                                                                                                                                                                                                                                                                    NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Virginia
Y: United
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2035 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                             uss: single
linear
                                                                                                                                                                  Peptide
147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IBM PC compatible SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                     Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
                                                                                                                                                                                                                                                                                                                                                                                                              128..1899
                                                                                                                                                                                                                                                          Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                 peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   States
 8.5%;
24.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US 08/039,320
                                                                                                                                                                                                       /note= "Amino acid 144 is Y =
which is the codon for Thr or
or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US/08/479,537A
                                                                                                                                                                                                                                                                                  /note= "Amino acid 134 is X1 = Xaa Xaa Xaa which is the codon for Pro or Ala wherein Pro = CCC, CCA, or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                               /note- "The amino acids spanning /note- "The amino acids spanning to 1899 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                           /note= "Amino acid 147 is X2 = Xaa which is the codon for Pro or Ala wherein or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FR 90/13101
                                                               /note= "Amino
21 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US 08/403,576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  017753-025
 Score
Pred.
  423
No.
                                                                precursor
                                                                            acids 1 to
  3.5; DB 2;
. 3.9e-19;
                                                                o 21 are a sequence.
                Length
                                                                                                                                                                                                               = Xaa
r Asn wherein
."
                                                                                                                                                                                                                        Thr
                                                                                                                                              Pro =
                                                                                                                                                                                                                           il
                                                                                                                                              CCT,
                                                                                                                                                                                                                          ACT,
                                                                                                                                                ccc,
                                                                                                                                                                                                                          ACC,
                                                                                                                                                                                                                                                                                                     CCT,
                                                                                                                                                CC
                                                                                                                                                                                                                           AC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δÃ
                                                                                                                                                                                                               망
                                                                                                                                                                                                                                      δõ
                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                              QΥ
                                                                                                          밁
                                                                                                                                 QΥ
                                                                                                                                                            B
                                                                                                                                                                                   δÃ
      RESULT 14
US-09-083-116-2
                                                        뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                                                                                                                             524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               469
                                                                                                                                                                                                                                                                                                                    645
                                                                                                                                                                                                                                                                                                                                           618
                                                                                                                                                                                                                                                                                                                                                                    586
                                                                                                                                                                                                                                                                                                                                                                                             572
                                                         925
                                                                                  883
                                                                                                           869
                                                                                                                                                             821
                                                                                                                                                                                     764
                                                                                                                                                                                                               763
                                                                                                                                                                                                                                                                 703
                                                                                                                                                                                                                                                                                          662
```

```
123 -- DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          134 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         531 TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS----APDXR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16 TVLTV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               75 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APXAHGYTSAPDXRPXPGSTAPXAHGYTSAPDXRP-XPGSTAPXAHGYTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         {\tt APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PKE--PAPTTKEP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TTPKEPAPTTT-----KSAPTTPKEP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTTSD
ŖΡ
                                              ΤP
                                                                                                                                             KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT-SRIAEAMLQTTTRPNQ
                                                                                                                                                                                                                                                     TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ
                                                                                                                                                                                                                                                                                                           DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                            -TKPEMTTTAKDK----TTERDLRTTPETTTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                                                                                                                                                                                                                                                                             {\tt DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTTKEPTT - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KPAP--TTPK----ETAPTTPKEPAPTTP-----KKPAPTTPETPPPTTSEVSTP 661
                                                                                               -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPXPGSTAPXAHGVTSAPDX
                                                                                                                                                                                                         APDXRPXPG----
926
                                              884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                      -STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -APTTTKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                                      820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            707
```

```
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION:
; OTHER INFORMATION:
US-09-083-116-5
                                                                                             Ş
                       밁
                                            Š
                                                                       В
                                                                                                                        밁
                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 0177/
REFERENCE/DOCKET NUMBER: 0177/
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEPHAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 anino acids
                                                                                                                                                                          Matches
                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME/KEY: Peptide
LOCATION: 128.1727
OTHER INFORMATION: /r
OTHER INFORMATION: 12
OTHER INFORMATION: 20
OTHER INFORMATION: re
FEATURE:
NAME/KEY: Peptide
                                                                                                                                                                                                                                                                                                         FEATURE:
                       123
                                               177
                                                                                              134
                                                                                                                                                                                                                                                                                                                                                                                   FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: 134
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: WO PERIOR DATE: 23-OCT-1991
PRIOR APPLICATION DATA:
                                                                       66
                                                                                                                                                                                                                                                                                                                 OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                        NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION: 144
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                         75 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKYLAKPT
                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
STRANDEDNESS: sir
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Teskin, Robin L. REGISTRATION NUMBER: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: US 0: FILING DATE: 04-APR-1993
                                                                                        P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP
                                                                                                                 TVLTV-----VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
              --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                           TTPKEPAPTTT----KSAPTTPKEP
                                                                 PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP
                                                                                                                                                                         235;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION DATA:
                                                                                                                                                                                     Similarity
                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                          Peptide
147
                                                                                                                                                                                                                                                                                                                                                                                                                                    Peptide
144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   single
                                                                                                                                                                                   8.5%;
                                                                                                                                                                                                                                                                                                                                                                                             /note= "A
which is
or ACG; a
                                                                                                                                                                                                                                                                                                          /note= "Amino acid 147 is X2 = Xaa
which is the codon for Pro or Ala wherein Pro =
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                  /note= "Amino acids 1 to 21 are a
21 amino acid precursor sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note= "Amino acid 134 is X1 = Xaa which is the codon for Pro or Ala wherein Pro or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "The amino acids spanning 128 to 1727 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US 08/403,576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US 08/039,320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO PCT/FR91/00835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35,030
                                                                                                                                                                        63;
                                                                                                                                                                     Score 423.5; DB 4;
Pred. No. 3.5e-19;
3; Mismatches 461;
                                                                                                                                                                                                                                                                                                                                                                                       "Amino acid 144 is Y = Xaa
is the codon for Thr or Asn wherein Thr =
; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        017753-025
                                        -APTITKEPAPTTPKEPAPTT
                                                                                                                                                                     Indels
                                                                                                                                                                                             Length
                                                                                                                                                                     203;
                                                                                                                                                                  Gaps
                                                                                                                                          133
                                       217
                                                                  122
                                                                                                                   65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                                                                                         CCT,
                                                                                                                                                                                                                                                                                                                                                                                                     ACT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CCT,
                                                                                                                                                                    48;
                                                                                                                                                                                                                                                                                                                         , ၁၁၁
                                                                                                                                                                                                                                                                                                                                                                                                     ACC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ccc,
                                                                                                                                                                                                                                                                                                                         င္ပ
                                                                                                                                                                                                                                                                                                                                                                                                      Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                င္ပ
                                                                                                                                                                            US-08-479-537A-2
                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                          QV
                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         문
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
STREET:
CITY: /
                                                                                                                                                                                                                                 925
                                                                                                                                                                                                                                                                                                                                     821
                                                                                                                                                                                                                                                                                                                                                              764
                                                                                                                                                                                                                                                                                                                                                                                       763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 662
                                                                                                                                                                                                                                                                                                                                                                                                                                        703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                       ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     218
                                                                                                                                                                                                                                 RΡ
                                                                                                                                                                                                                                                           ŢΡ
                                                                                                                                                                                                                                                         884
Alexandria
```

```
Sequence 2, Application US/08479537;
Patent No. 5861381
GENERAL INFORMATION:
APPLICANT: CHAMBON, Pierre
APPLICANT: KIENY, Marie-Paule
APPLICANT: LATHE, Richard
APPLICANT: HARBUVENI, Marie
                                               TITLE OF INVENTION: PH
TITLE OF INVENTION: TR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            708 -TKPEMTTTAKDK----TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             618 KPAP--TTPK----ETAPTTPKEPAPTTP-----KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  586 PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               572 KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  531 TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS----APDXR 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472 APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469 TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT-SRIAEAMLQTTTRPNQ 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPXPGSTAPXAHGVTSAPDX 924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APDXRPXPG----STAPXAHGVT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ 823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTP-----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP 571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAPTTPKEPAPTTPK-----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TKEPSPTTPKEPAPTT--TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKE--PAPTTKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                   Application US/08479537A
E: BURNS, DOAL P.O. Box 1404
                                                                               TREATMENT
5
                                                                                            PHARMACEUTICAL COMPOSITION TREATMENT OR PREVENTION OF
                       DOANE, SWECKER & MATHIS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                       L.L.P.
                                                                                            FOR THE
A MALIGNANT
                                                                                               TUMOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 264
```

```
рь
                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                   Q
          Оy
                                    DЪ
                                                           Š
                                                                                       В
                                                                                                           δÃ
                                                                                                                                       멍
                                                                                                                                                              δÃ
                                                                                                                                                                                          В
                                                                                                                                                                                                               VQ
                                                                                                                                                                                                                                                                                             B
                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
Matches 235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: (703) 836-202
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELECOMMUNICATION INFORMATION: TELEPHONE: (703) 836-6620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 1867 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURE:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME/KEY:
LOCATION:
                                                                                                                                          181
                                                                                                                                                                                            123
                                                                                                                                                                                                                     177
             361
                                      301
                                                                                        241
                                                                                                                                                                    218
                                                                                                                                                                                                                                                                                                                                              Local Similarity
hes 235; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
                                                                                                                                                                                                                                                                                               16
                                                                                                                                                                                                                                                                                                              75 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                             66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                   P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP 176
                                                                                                                                                                                                                                                                                             TVLTV-----VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT 412
                                      APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 359
                                                                                          APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                          APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                  TKE--PAPTTT----
                                                                                                                                                                                            --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                    TTPKEPAPTTT-----KSAPTTPKEP-----APTTTKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                             PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP
                                                                TKEPSPTTPKEPAPTT - - TKSAPTTTKEPAPTT - - - -
                                                                                                                PKE--PAPTTKEP-----APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide
134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
128..1727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3) by

) 836-2021

-- NO: 5:
                                                                                                                                                                                                                                                                                                                                                            8.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note= "The amino acids spanning 128 to 1727 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "A
21 amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Au
which is
or CCG; a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "Amino acid 144 is Y = Xaa
which is the codon for Thr or Asn
or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Amino acid 147
which is the codon for
or CCG; and Ala = GCT,
                                                                                                                                                                    -KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                  63;
                                                                                                                                                                                                                                                                                                                                                Score 423.5; DB 2;
Pred. No. 3.5e-19;
3; Mismatches 461;
                                                                                                                                                                                                                                                                                                                                                                                                                           "Amino
no acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Amino acid 134 is the codon for and Ala = GCT,
                                                                                                                                                                                                                                                                                                                                                                                                                          ) acids 1 to
1 precursor s
                                                                   -TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         is X2 = Xaa
Pro or Ala wherein Pro
GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             is X1 = Xaa
Pro or Ala wherein Pro =
GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                        21
                                                                                                                                                                                                                                                                                                                                                                                                                                        are
                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                         Length 1867;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   wherein
                                                                                                                                                                                                                                                                                                                                                    203;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Thr
                                                                                                                      311
                                                                                                                                                                        264
                                                                                                                                                                                                180
                                                                                                                                                                                                                         217
                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                   65
                                                                                                                                                                                                                                                                                                                           133
                                                                                            300
                                                                                                                                              240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CCI, CCC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCI, CCC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT,
                                                                                                                                                                                                                                                                                                                                                     48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ф
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-083-116-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 5, Application US/09083116 Patent No. 6203795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                             APPLICANT: LATHE, Ric
APPLICANT: HARBUVENI
TITLE OF INVENTION: I
TITLE OF INVENTION: I
TITLE OF SEQUENCES:
                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
                  PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             413 TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   360 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             925 RP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   883 TP
                                                                                                                                                                                                                                                          CITY: Alexandria
STATE: Virginia
                                   CLASSIFICATION:
                                                        FILING DATE:
                                                                             APPLICATION NUMBER:
                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                  STREET:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                       ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX 644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS----APDXR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TPKA----APVITEKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP 523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -TKPEMTTTAKDK---TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS 530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT-SRIAEAMLQTTTRPNQ 882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APDXRPXPG-----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KPAP--TTPK----ETAPTTPKEPAPTTP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPXPGSTAPXAHGVTSAPDX 924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     884
                                                                                                                                                                                                                    : Virginia RY: United States 22313-1404
                                                                                                                                                                                                                                                                                                       E: BURNS, DOA
P.O. Box 1404
                                                                                                                                                                                                                                                                                                                                                                                                                           HAREUVENI, Mara
                                                                                                                                                                                                                                                                                                                                                                                                                                           KIENY, Marie-Paule
LATHE, Richard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAMBON,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP 571
                                                                                                                                                                                                                                                                                                                                                             PHARMACEUTICAL COMPOSITION FOR THE TREATMENT OR PREVENTION OF A MALIGNANT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pierre
                                                                                                                                                                                                                                                                                                                          DOANE,
                                                                                US/09/083,116
  08/479,537
                                                                                                                                                                                                                                                                                                                            SWECKER
                                                                                                                                                                                                                                                                                                                            Β'n
                                                                                                                          Version
                                                                                                                                                                                                                                                                                                                            MATHIS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                           TUMOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  707
```

```
밁
                                          Ş
                                                                                                            Ş
                                                                              밁
                                                                                                                                                  밁
                                                                                                                                                                            Ş
                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                 á
                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                     : OTHER INFORMATION: Description of Artificial Sequence: Illustrative : OTHER INFORMATION: amino acid sequence US-09-219-849-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
US-09-219-849-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION NUMBER: US/09/219,849
CURRENT FILING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PATENTIN Ver. 2.1
SEQ ID NO 5
LENGTH ACC
                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 173; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: WIND, RICHELE D.

APPLICANT: VAN DEN BOSCH, TANJA J.

TITLE OF INVENTION: SITVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUTTABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
FILE REFERENCE: 2728-2

CURRENUM NOTICE: 2728-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 5, Application US/09219849 Patent No. 6150081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH: 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Artificial Sequence FEATURE:
                               331 APTITKEPAPTITKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTP---KKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1198 QOTDEITGKPVDTVTGLPYDPSTGEIIDP 1226
                                                                                                                                                  187
                                                                                                                                                                                                                      127
                                                                                                                                                                                                                                        190 APTTPKEPAPTTTKEPA-PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP---KKPA 245
                                                                                                                                                                                   246
                                                                                                                                                                                                                                                                                                              132 PTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAP--TTTKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1144 KDGLIVPPTNSINKDPVTNTQYSNTTGNI----INP--ETGKVIPGSLPGSLNYPSFNTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1024 VYGLPVGEITGLPKDPGSDIPFNSTTGELVDPSTGKPINNSTAGIVSGKPGLPPIEDENG 1083
                                                                                                                                                                                                                                                                           72 PGPAGPPGSRDPG--PPGAPGPAGP--PGSRDPGPPGAPG-PAGPPGSRDPGPPGAPGPA 126
                                                                                                                                                                                                       GPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPPGSRDPGPPGAPGPAGPPGSRDPG 186
         GPPGSRDPGP----
                                                                   PAGPKGAHGPAGPKGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGP---PGAPGP----A
                                                                                                                                     P--PGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAHGPAGPKGAHGPAGPKGAHG
                                                                                                                                                                     PTTPKEPAPTTP----KEPTPTTPKEPA---PTTKEPAPTTPKEPA--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EVTPD------MDYLPRVPNQGIIINP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----LNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NLFDPSTNLPIDGNNQLVNPETNSTVSGSTSGTTKPKPGIPVNGGGVVPDEEAKDQADKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IMNKP--EETAKPKD-----RATNSKATTPKPQKPTKAPKKPTSTKKPKTMP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----RVRKPKTTPTPRKMTS--TMPE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VAN HEERDE, GEORGE V
VAN RIJN, ALEXIS C.
BOUWSTRA, JAN B.
DE WOLF, FREDERIK A.
                                                                                           ----PTAPK-KPAPTTP---KEPAPTTPKEPA-PTTTKEPSPTTPKEPAPTTTKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DE WOLF, FREDERIK A.
MOOBROEK, ANDREAS
WERTEN, MARC W.T.
                                                                                                                                                                                                                                                                                                                                                        Conservative
---PGAPGPAGPPGSRDPGP--PGAPGPAGPPGSRDPGP--PGAPGP 348
                                                                                                                                                                                                                                                                                                                                                                         9.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GEORGE V.
                                                                                                                                                                                                                                                                                                                                                        60;
                                                                                                                                                                                                                                                                                                                                                     Score 452; DB 4;
Pred. No. 2.8e-21;
50; Mismatches 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICATION AND ALSO THE
                                                                                                                                                                                                                                                                                                                                                                                    Length 960;
                                                                                                                                                                                                                                                                                                                                                   Indels 134;
                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                     298
                                                                                                                                       244
                                                                                                                                                                          284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            842
                                                                                                                                                                                                                                                                                                                                                 41;
```

```
RESULT 11
US-08-479-537A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5, Applic Patent No. 5861381
                                                                                                                                                          FILING DATE: 23-OCT-19
PRIOR APPLICATION DATA:
APPLICATION NUMBER: W
FILING DATE: 23-OCT-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION
                                              APPLICATION NUMBER: US 08/403,576 FILING DATE: 14-MAR-1995 ATTORNEY/AGENT INFORMATION:
                                                                                            APPLICATION NUMBER: US 0 FILING DATE: 04-APR-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                               PRIOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,537A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: LATHE, RI
APPLICANT: HAREUVENI
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
               NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          696 SRDPGPPGAHGPAGPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            694
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              523 PPGSRDPGPPGA--PGPAGPPGSRDPGPPGAPGPPGSRDPGPPGAPGPAGPPGSRDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                              CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        463
                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAPTTP----ETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALEN-----S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GP--PGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGP--PGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APTTPKKPAP-----KELAPTTTKGPTST--TSDKPAPTTPKETAPTTPKEPAPTTPKK 642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PKEPGVPTTKTPAATK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PGPAGPPGSRDPGPPGAPGPAGPPGSRDPGP-PGAPGPAGPPGSRDPGPPGAPGPAGPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GAHGPAGPKGAHGPAGPKGAHGPAGPKGAHGPAGPKGAPGPAGPPGSRDPGPPGAPGPAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPKEPA-----PTTPKET-APTTPKGT-----APTTLKEPAPTTPKKP-APKELA 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SRDPGPPGAPGPAGPPGSRDPGP--PGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTTTKEPTSTTSDKPAPTTPKGTA-PTTPKEPAPTTP---KEPAPTTPKGTA-PTTLKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PEKPAPT-TPEELAPTTPEEPTPTTPEEPAPTTPKAAA-PNTPKEPAPTTP---KEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGP----PGAPGPAGPPG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TTP---KEPAPTTTKKPA-----PTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Application US/08479537A
                                                                                                                                                                                                                                                                                                                                                                                                                                             Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P.O. Box 1404
                                                                                                                                                                                                                                                                                                                                                                                                                              United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HAREUVENI, Mara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KIENY, Marie-Paule
LATHE, Richard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAMBON,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BURNS,
                                                                                                                                                               23-OCT-1991
                                                                                                                                                                                                              23-OCT-1990
   NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHARMACEUTICAL COMPOSITION TREATMENT OR PREVENTION OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOANE, SWECKER & MATHIS, L.L.P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pierre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         709
                                                                                                                               US 08/039,320
                                                                                                                                                                             WO PCT/FR91/00835
                                                                                                                                                                                                                             FR 90/13101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           711
017753-025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FOR THE
A MALIGNANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             462
```

RESUUS-0	0
SULT 9 S-08-3618-6 S-08-08-3618-6 S-08-08-08-08-08-08-08-08-08-08-08-08-08-	

по	п ю	U 0	a a	D 0	<u>ي</u> و	р. ю	<u>p</u> .0	B 8	D 5	g p	Q	DŁ Qy	Qy dd	Qy Db	Оу Оъ	Qу Дъ	Qу Дъ	
Qy 739 TKETATTTEKTIESKITATT-TQVTSTTTQDTTPFKITTLAPKVTTTKKTITTE 797 :	KM 73 NP 96	9 6	 SALIDPA	/1 PKEPARTIPAGIAFILAR FREITING FOR INNIE I	SI PRI PER LINDERS SI PRIGDIEITPIPIEKMIDKYTRMIYDYNSGLLLDSNDEPIPGSQAGQIADTSNLFPVQTH SI PRI PRI PRI PRI PRI PRI PRI PRI PRI PR	471 KAAADNIPKEPARTITEKEEARTITEKEEARTITEKEEARTITETUPTTTITTKRDEMITTITEL 68 628KVPGKPPIATTTTIKPIVTITTTKATTTTTTVPTTTITTKRDEMITTTTPL 68	PTEKEKAPITENALIFITEDALOFITEDAK FARIA 1 :	369 PAPTTPKKPAPTTPKE	67 VKDVGNTISVRCRKGAGKLEFPDRSLDFTIPPVAGHNSCSIIVGVSGDGKIHVSPYGSKD	407 TSETESVIKPDEWCWLEKNGECEAKGATYVGVIGNUGKLENGMAFIMIENUULITVALAKA 190	350 PKE	QY 290 KPAPTTPKEPAPTTPKEPAPTTTKEAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTKSAPTTTTKAPTTTTTTTTTT	' ×	SA 2 TT 3	110 ETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTP 1	65 LPPNSDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTTSAK	5 KKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTOHNKVSTSPKITTAKPINPRPS :	Query Match 9.2%; Score 463.5; DB 3; Length 1721; Best Local Similarity 22.1%; Pred. No. 9.7e-22; Matches 271; Conservative 78; Mismatches 467; Indels 413; Gaps
																		39

Indels 507; Length 1721

Gaps

45;

249 92

```
APPLICANT: PEERSEM, CAROLYN

APPLICANT: LEECH, JAMES

APPLICANT: NELSON, RICHARD, C.

APPLICANT: NELSON, RICHARD, C.

APPLICANT: OTIVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS /

TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF CRYPTOSPORTIGHUM PAIN

TITLE OF INVENTION: INFECTIONS

FILE REFERENCE: 480.19-4(HV)

CURRENT APPLICATION NUMBER: US/08/700,651B

CURRENT FILING DATE: 1997-08-14

EARLIER APPLICATION NUMBER: 08/415,751

EARLIER APPLICATION NUMBER: 08/415,751

EARLIER FILING DATE: 1995-04-03

NUMBER OF SEQ ID NOS: 15

COPPUANDE: DESCRIPTION OF CRYPTOSPORTER

ANDIBLE OF SEQ ID NOS: 15
                                                                                                                                                                                                                                                                                        RESULT 8
US-08-700-651-5
I Sequence 5, Application US/08700651B
Patent No. 6015882
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             당
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                     SOFTWARE: PatentIn Ver. SEQ ID NO 5
                                         ENGTH: 1721
                                                                                                                                                                                                                                                                                                                                                                                                                                          712 MTTTAKDKTTERDLRTTPETTTAAP 736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       253 AKPTYKAKPTYKAKPTYPSTYKAKPTYP----PTYKAKPSYPPTYKAKPTYKAKPTYKAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           330 SAPTTIKEPA-----PTITKSAPTIPKEPSPTITKEPA---PTIPKEPA----PTIPKK 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197 YP----PTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPSLIKAKPSYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                           YPPTHKAKPTYKAKPSYPPTHKAKP 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYK-AKPTYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKAKPTYPSTYKAKPSYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PA--PTTPKEPA--PTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEK 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KPTYPSTYKAKPTYK--AKPSYP----PTYKAKPS----YPPTYKAKPSYPPTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA--PTTPKKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPTYPSTYKAKPSYPPTYKPKISYPPTYKAKPSYPSTYKAKSSYPPTYKAKPSYPPTYKA 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AKPTYPSTYKAKPSYP----PTYKAKPSYPPTYKAKPSYPPTYK-AKPSYP----PTYKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LAPTTPE--KPAPTTPEELAPTTPEEPT--PTTPEEPA-PTTPKAAAPNTPKEPAPTTPK 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTYKAKPSYPPT----YKAKPSYPPTYKAKPTYKA-----KPTYPST---YKAKPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPA--PTTPKKPA-PTT----PETP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKVKPT--YK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TPKEPAPTAPKKPA--PTTPKEPA--PTTPKEPAPTTTKEPSPTTP----KEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTTKSAPTTPKEPAPTTPKKPAPT-TPKEPAPTT----PKEPTPTTPKEPAPTTKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTYKPKPSYP--PSYKTKKTYPPTYKPKLTYPPTYKPKPSYPPSYKPNKTYPPTYKPKLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYPPTYKSKPTYKPKITYPPTYKAKPSYPPTYKPKKTYPPT-YKPKLTYPPTYKPKASYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTPTTIKSAPT-TPKEPAPTTIKSAPTTPKEPAPTTTKEPAPTTTKEPA--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKE----PASTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PKMTYPPTYKPKPSYPP------TYKSKPTY-----KPKTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -----YPPTYK--AKPS------YPPTYKPKKTYPPTYKPKLTYPPTYKPKP
                                                                                                                                                                                          , DNAS AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79
                                                                                                                                                                                                        AND
                                                                                                                                                                                                         RNAS
                                                                      Q
                                                                                                        Ъ
                                                                                                                                   Qy
                                                                                                                                                                    밁
                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                     DЬ
                                                                                                                                                                                                                                                                  ρ
                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Cryptosporidium parvum US-08-700-651-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 9.3
Best Local Similarity 20.9
Matches 285; Conservative
                                         1074
                                                                                                      1014
           633
                                                                                                                                       551
                                                                                                                                                                                                    497
                                                                                                                                                                                                                                     899
                                                                                                                                                                                                                                                                     469
                                                                                                                                                                                                                                                                                                  839
                                                                                                                                                                                                                                                                                                                                   440
                                                                                                                                                                                                                                                                                                                                                                  779
                                                                                                                                                                                                                                                                                                                                                                                                                               719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              544 YCDTCTAKYGAIHSGYQTSADFYTTTTAKPTTTTTGAPGQPTTTTTGSPSKPTTTTTKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484 KLEFPDRSLDFTIPPVAGHNSCSIIVGVSGDGKIHVSPYGSKDVSLISAPIOPSELFNEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         424 KNGECEAKGATYVGVIGKDGRIENGMAFTMIPNDDTHVRFRFKVKDVGNTISVRCRKGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVSNGIKTDNVYGLÞVGEITGLÞKDÞGSDIÞFNSTTGELVDÞSTGKÞINNSTAGIVSGKÞ
                                                                                                                                                                                     ----PAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT--TTKEPTSTTSDKP 550
          KEPA --
                             GLPPIEDENGNLFDPSTNLPIDGNNQLVNPETNSTVSGSTSGTT-KPKPGIPVNGGGVVP
                                                                       GTAPT---
                                                                                                                                                        GDLNPIMKTPTQTDSVTGKPIDPTT---GLPFNP--PTGHLINPTNNNTMDSSFAGAYKY 1013
                                                                                                                                                                                                                                                        APTT----
                                                                                                                                                                                                                       PPEVAAANADKFKLSIPPSVPESIPEKDQKIDSISELMYDIESGRLIGQVSKRPIPGSIA
                                                                                                                                                                                                                                                                                     ISASESILSQKSALIDPATNMYVGEFGGLLNPATGVMIPGFLGPSEQTQFSPEIEDGGII
                                                                                                                                                                                                                                                                                                                                                    LAAKNLTVDTDETYGLPIDTLTGYPLDPVSLIPFNPETGELFDPISDEIMNGTIAGIVSG
                                                                                                                                                                                                                                                                                                                                                                                      PAP-----TTPKETA-PTTP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               T----TTTTKRDEMTTTTPLPDIGDIEITPIPIEKMLDKYTRMIYDYNSGLLLDSNDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTTTTTLNPIITTTQKPTTTTTTKVPGKPPIATTTTLKPIVTTTTKATTTTTTVPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITKSA-----PTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPT 380
                                                                                                                                                                                                                                                                                                                       PIPGSQAGQIADTSNLFPVQTHKSTGLPIDPMVGLPFDPKSGNLVHPYTNQTMSGLSVSY 778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPKEPAPTTPKEPAPTTTKKPAP------TAPKE 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAPTTTKEPAPTTTKSAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTTTTTTTTKPTTTTTTTT-----TTTKKPTTTTTATTTTTSETESVIKPDEWCWLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPT 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --TLKEPAPTTPKKPAPKELAP---TTTKGPTSTTSDKPAPTTPKETAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----PTTPKEPAPTT---PKEPAPTTTKEPS-PTTPKEPAP
                                                                                                                         ----PKGTAPTTPKEPA-----PTTPKEPAPTTPK-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9.3%; Score 466.5; DB 3
20.9%; Pred. No. 6.3e-22;
-PTTPKKPAPTTPETPPPTTSEVSTPTTTKE-PTTIHKS---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100; Mismatches 473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----EPAPTTPKEPA-PTAPKKPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 3;
```

1073 632

468

439

718

663

603

543

292

423

267

273

369

212

밁 Qy

444 PNKPNP----NEPSNPNKPNPN-----EPLNPNEPSNPNEPSNPNAPSNPNE--PSNPNE 492

```
QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Вp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            맑
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            γo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 6
PCT-US92-00018-2
PCT-US92-00018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 2, Application PC/TUS9200018 GENERAL INFORMATION:
                                        TELEFAX: (301) 295-4033
INFORMATION FOR SEO ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
TYPE: AMINO ACID
                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Khusmith, Srisin
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective management of INVENTION: immunogen and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  393
                                                                                                                                NAME: Spevack, Aviam D. TELECOMMUNICATION INFORMATION: TELEPHONE: (301) 295-6759
                                                                                                                                                                                                                                                       SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   631 TPKEP-----APTTPKKPAPTTPETPPPTTSEVSTPTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         571
                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  707
                  MOLECULE TYPE:
                                                                                                                                                                                                         APPLICATION NUMBER: FILING DATE: 199201 CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
ZIP: 20814-5044
                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: A. David Spevack
STREET: NMRDC Building 1 T-12 National Naval
STREET: Medical Center
                                TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SNPNE--PLNPNEP-----SNPNEPSNPNEPSNPEE--PSNPKE--PSNPNE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTPEEPTPTTPEEPAPTTPKAAAPNTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAPTTTKKPAPTAPKEP-APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAP 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LPIIPQKGNNIPSNLPENPSDSEVEYPRPNDNGENSNNTMKSKKNI---PNEPIPSPGDN 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------PSNPEEPNPEE--PSNPKEP----SNPEEPINPEELNPKEPSNPEESN 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PYKGHEERIPKPHRSNDDYVYDNNVNKNNKDEPEIPNNE 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKEPI-----NPEESNPKEP----INPEDNENPLIQDEPIEPRNDSNVIPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTPKALENSPK - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                      Bethesda
                                                                                                                                                                                                                                                                                                                                                                                                    ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hedstrom, Richard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hoffman, Stephen L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Charoenvit,
                                                                                                                                                                                                                           19920103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            immunogen and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protective malaria sporozoite surface
                                                                                                                                                                                                                                             PCT/US92/00018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EPGVPTTK 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         564
```

QΥ

Indels 169;

Gaps

45;

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Вp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
5202236-13
;Patent No. 5202236
                                                                                             5202236-13
                                                                                                                                                                                                                                                                                                                                                    ; APPLICANT: MAUGH, KATHY J.;ANDERSON, DAVID M.;STRAUSBERG;SUSAN L.;MCCANDLISS, RUSS;WEI, TENA;FILPULA, DAVID TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
                                                                                                                                                                                                                                                                                                                                      PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                SEQ ID NO:13:
Query Match
Best Local Similarity
Matches 221; Conserv
                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES: 39 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      168 TPTTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTTKE 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    571 PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           650 LPIIPQKGNNIPSNLPENPSDSEVEYPRPNDNGENSNNTMKSKKNI---PNEPIPSPGDN 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     607
                                                                                                                                                                       APPLICATION NUMBER: 82,456
RILING DATE: 07-AUG-1987
RPPLICATION NUMBER: 933,945
RILING DATE: 24 NOV-1986
APPLICATION NUMBER: 650,128
                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/07/528,762 FILING DATE: 25-MAY-1990
                                                                                                                                                                     APPLICATION NUMBER:
                                                                                                            LENGTH: 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TKEPAPITIKSAPITPKEPSPITIKEPAPITPKEPA-PITPKKP-APITPKEPAPITPKE 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAPTTTKSAPTTPKEP-APTTPKKP-APTTPKEP-APTTPKEP-TPTTPKEPAPTTKEPA 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SNPNE--PLNPNEP----SNPNEPSNPNEPSNPEE--PSNPKE--PSNPNE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPTTTKKPAPTAPKEP-APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAP 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PNKPNP----NEPSNPNKPNPN-----EPLNPNEPSNPNEPSNPNAPSNPNE--PSNPNE 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RRNPKRRNPNKPNKPNKPNPNEPSNPNKPNPN----EPSNPNKPNP----NEPSN 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTTPKEPAPTAPK--KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTT 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPKEP-----APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TIPEEPTPTTPEEPAPTTPKAAAPNTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PS-----NPNEPSNPNEPSNPNE--PSNPKK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PYKGHEERIPKPHRSNDDYVYDNNVNKNNKDEPEIPNNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKEPI---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTPKALENSPK-----
 9.5%; Score 476.5; DB 6;
ilarity 29.7%; Pred. No. 5.3e-23;
Conservative 71; Mismatches 284;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----PSNPEEPNPEE---PSNPKEP----SNPEEPINPEELNPKEPSNPEESN 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----NPEESNPKEP----INPEDNENPLIIQDEPIEPRNDSNVIPI 649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9.7%; Score 488.5; DB 5; 26.4%; Pred. No. 1.2e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EPGVPTTK 703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------PSNPNE--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 826;
                                          Length
                                                  652;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34;
```

0 9 O D O

```
δÃ
                 밁
                                                                                             В
                                                                                                                                Qy
                                                                                                                                                                          밁
                                                                                                                                                                                                             ρy
                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE: US-07-638-431-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULY 5
US-07-638-431-2
; Sequence 2, Application US/07638431
; Patent No. 5198535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                          Query Match 9.7%; Score 488.5; DB 1; Best Local Similarity 26.4%; Pred. No. 1.2e-23; Matches 153; Conservative 54; Mismatches 223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ZIP: 20814-5044

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA;
                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATTORNEY/AGENT INFORMATION:
NAME: Spevack, Avrom D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
393 RRNPKRRNPNKPKPNKPNPNKPNPNEPSNPNKPNPN----EPSNPNKPNP----NEPSN 443
                                      277 PTTPKEPAPTAPK--KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTT 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: A. David Spevack
STREET: NMRDC Building 1 T-12 National Naval
STREET: Medical Center
CITY: Bethesda
STATE: MD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria sporozolte surface protein
TITLE OF INVENTION: immunogen and gene
                                                                                             333
                                                                                                                                    221
                                                                                                                                                           273 TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                                                                168 TPTTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTTKE 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Hoffman, Stephen L
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard
APPLICANT: Khusmith, Srisin
APPLICANT: Rogers IV, William
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1466
                                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 826 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                850 TPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPK-SEDA 898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        739 TKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US
FILING DATE: 19910110
CLASSIFICATION: 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                               PAPTITKSAPTTPKEP-APTTPKKP-APTTPKEP-APTTPKEP-TPTTPKEPAPTTKEPA 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GFINGTSGEQSHEKDPSTG-----KPLDPNTGLPFDEDSGSLINPETGDKLQGSHSGTF 1465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MPVPGKPQGENGGIMTPEQILEAL-----NKLPTSNEVNISPRPSSDA 1508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --IMNKPEETAKPKDRATNSKATTPKPQKP-TKAP---KKPTSTKKPKTMPRVRKPKT--- 849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DLETGLPRDP-----VSGLPQLPNGTLVD-------PSNKKPIPGSHS 1411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Stephen L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US/07/638,431
                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                   Length 826;
                                                                                                                                                                                                                                          149;
                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            797
                                                                                                                                                                                                                                          34;
```

```
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-103-429A-3
                                                                                                                                                                                                                                                                                                    Вp
                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                               Qy
                                                                                                                                                                                 멍
                                                                                                                                                                                                                                    δÃ
Ş
                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GENERAL INFORMATION APPLICANT: Granac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: (607) 256-36 (NFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE OF INVENTION: Å No. 6187558el Invertebrate Intestinal Mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
LENGTH: 786 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A
REGISTRATION NUMBER: 34,390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE/DOCKET NUMBER: BT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (607) 256-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              132 PTPKAETTTKGPALTTPKEPTPTTTKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAP 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC
OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: 118 |
CITY: Ithaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  || : ||||: ||||: ||||: ||||: ||||: ||||: |||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  192 TTPKEPAPTTTKEPAPTTPKEPAPTTTKSAPTTPKEPAPTTPKE 251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: FILING DATE: 24-JUN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
      362
                                                                                                                                                                                                                                                                                                    242 ELLPNGCPADFDIHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVEC 301
                                                                                                                                                                                                                                                                                                                                                                                                                         184 -APTTTQAPTTITQAATTPAATTPAATTPAATTPAA-TTPAATTPGVPAPTSAPVWPPIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      252 PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTT-PKEPA-- 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Trichoplusia ni
TISSUE TYPE: peritrophic membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     93 PGPPAE-TTQAPATT----QAPTTTQAPTTTT------QAPTTTTQ--ATTTTQAP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS:
      APTT----PKEP--
                                                            CGAGTHFSFELQQCDHIELVGCTLPGGESEEVDVDEDA-CTGWYCPTEPIEWEPLPNGCP
                                                                                                                                                                                 \tt DGEISPAPPVTEGNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERR
                                                                                                                                                                                                                                              ----EPAPTTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ΥN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: Brown, Pinnisi & Michaels, P.C.
118 No. 6187558th Tioga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wang, Ping
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             786 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Granados,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E: Floppy disk
IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PatentIn Release #1.0, Version
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ĕ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-JUN-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          256-3628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Robert R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US/09/103,429A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ω
••
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BTI-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 498.5; DB 4;
Pred. No. 2.6e-24;
                                                                                                                               KSAPTTKEPAPTTKSAPTTPKEPSPTTKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 #1.30
                                                                                                                                                                                                                                                                                                                                                                                ---PTTTKEPSPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                               420
                                                                                                                                        361
                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                   321
                         381
                                                                                                                                                                                                                                                                 328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32;
```

```
QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   da
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
US-08-928-361B-5
                                                                  ; MOLECULE TYPE: US-08-928-361B-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 5, Application US/08928361B Patent No. 6071518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GENERAL INFORMATION:
Query Match
Best Local Similarity
                                                                                                                                                                                       TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: PC-POS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   421 ADFSIDHLLPHESDCGQYLQCVHGQTIARPCPGNLHFSPATQSCESPVTAGCQVFECDSD
                                                                                                                                                                                                                        TELECOMMUNICATION INFORMATION: TELEPHONE: 650-324-1677
                                                                                                                                                                                                                                                                          FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana
REGISTRATION NUMBER: 30,
                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 588
                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 1837 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             678 ---APATTPEDDDIDP--PLPNDPINP 699
                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 1:
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZIP: 94306-1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ADDRESSEE:
                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US 60/026,062
                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
                                                                                                      TOPOLOGY:
                                                                                                                     STRANDEDNESS
                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --PAPTAAPTAAPTTAAPESPTTVTVP-PTAAPTAAPTTAVPEIPITVT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EKPAPTTPEELAPTTPEEPTPTTPEEBAPTTPKAAAPNT--PKEPAPTTPKE 496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTP 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DKPAPTTPKETAPTTPKEPAPTTPKKP 643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TAAPTT-AAPAPNTTVTVPPTAAPTAAPPTVAH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTS 616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAPTAAPTAAPTTAVPEIPTTVTSPPTAAPTTAAP---APNTT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPK 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NQCTSTAAPTAAPTAAPTAAPTAAPTAAPTA--APSTVVPPATPPATAAPVPPTT---AIPT-- 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Palo Alto
                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      385 Sherman
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Petersen,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                             650-324-1678
                                                                                                                                                                                                                                                                                                                                                                                              JMBER: US/08/928,361B
12-SEP-1997
                                                                                                                                                                                                                                                                                                                               13-SEP-1996
                                                                                    protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Carolyn
PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
SPECIES INFECTIONS
   9.8%;
21.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERNY, JONES & BIKSA
n Avenue, Suite 6
                                                                                                                                                                                                                                                                       30,518
                                                                                                                                                                                           ū
                                                                                                                                                                                                                                                               480.76-1(HV)
     Score
Pred.
     493; DB 3;
No. 1.4e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Version
                       Length 1837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -APNTTAAPVTTTS 677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---SAPTAAPT 587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----VTVPP 632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANALOGS AND FRAGMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  480
```

50;

```
RESULT 2
$202236-25
;Patent No. 5202236
; Patent No. 5202236
; APPLICANT: MAUGH, KATHY J.;ANDERSON, DAVID M.;STRAUSBERG,
;SUSAN L.;MCCANDLISS, RUSS;WEI, TENA;FILPULA, DAVID
; TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE
          NUMBER OF SEQUENCES: 39
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/528,762
FILING DATE: 25-MAY-1990
APPLICATION NUMBER: 82,456
FILING DATE: 07-AUG-1987
APPLICATION NUMBER: 933,945
                                                                                                                                                                                                                                                                           686 NTTAAP-----PTTTSAPATT-PE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ELSAEPTPKALENSPKEPGVPTTKTPAATKPE
                                                                                                                                                                                                                                                                                                                                      AP----NTTVTAPPTAAPTT-AAPAPNTTVTVPPTAAPTAAP----
                                                                                                                                                                                                                                                                                                                                                         APTTPKETAPTTPKEPAPTTPEKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTP
                                                                                                                                                                                                                                                                                                                                                                                                      TTAVPEIPTT-----VTSPPTAAPTT-AAPAPNT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                 TTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTST-TSDKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PESPTTVTVPPTAAPTA----APTT----AVPEIPITVTSAPTAAPT--AAPTAAPTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ETAPTT----PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TAAPTAAPTAAPTAAPSTVVPPA-TPPATAAPVPPTTAIPT----PAPTAAPTAAPTTAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QCDHIELVGCTLPGGESEEVDVDEDACTGWYCPTEPIEWEPLPNGCPADFSIDHLLPHES 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERRCGAGTHFSFELQ 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVECDGEISPAPPVTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTPEELAPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DCGQYLQCVHGQTIARPCPGNLHFSPATQSCESPVTAGCQVFECDSDNQCTSTAAPTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTPAATTPAAT-TPAATTPAATTPAATTPGVPAPTS----APVWPPICELLPNGCPADFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QAPTTTTQATTTT------QAPTTTQAPTTT----QATTTQAPTTTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTKEPAPTTPKEPAPTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RDCAPGTEFKFSAQTCVHAALAGCTLPGPPAETT---QAPATTQ--APTTTQ--APTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TALGLVAARPEVSDAEKNPALHEPHPDXP---PAEQXXLLPXEYDCTKFYYCEYGLKFIAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAP-----TTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ------APTTPKKLTPTT------PEKLAPTTPEKPAP 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----EPAPTTTK 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PTAP--KEPAPT-----TPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               211;
                                                                                                                                                                                                                                                                                                                                                                                                      -TVTVPPTAAPTTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          342;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                   679
                                                                                                                                                                                                                                                                                                                                                                                                    641
                                                                                                                                                                                                                                                                                                                                                                                                                                   619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           227
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40;
                                                                                                                                                                                                                                                                                                                         Qy
 US-09-103-429A-3; Sequence 3, App.; Patent No. 618'
                                                                                                                                                                                                Ωy
                                                                                                                                                                                                                                 B
                                                                                                                                                                                                                                                             δÃ
                                                                                                                                                                                                                                                                                               Дb
                                                                                                                                                                                                                                                                                                                                                              Ъ
                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                              멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; LENGTH: 744
5202236-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 24-NOV-1986
APPLICATION NUMBER: 650,128
FILING DATE: 13-SEP-1984
SEQ ID NO:25:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 239
                                                                                             720 YKAKPTYPSTYKAKPTYPST
                                                                                                                                                                                                                                                                                                 547
                                                                                                                                                                                                                                                                                                                                592
                                                                                                                                                                                                                                                                                                                                                                499
                                                                                                                                                                                                                                                                                                                                                                                                                                439
                                                                                                                                                                                                                                                                                                                                                                                                                                                               479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              323
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24 PKMTYPPTYKPKPSYPP------TYKSKPTY-----KPKIT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        51 PKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKOTSTDGKEKTTSAKE 110
ce 3, Application US/09103429A
No. 6187558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PEKLAPTTPE--KPAPTTPEELAPTTPEEPT--PTTPEEPA--TTPKA--AAPNT----P
                                                                                                                                                                                                                     YKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAK
                                                                                                                                                                                                                                        ---TPETPPPTTSEVSTPTTTKE----PTTIHKSPD-----ESTPELSAEPT-PKALENS
                                                                                                                                                                                                                                                                                                                  -PTTPKKPAPKELAPTTTKGPT--STTSDKPAPTTPKETAPTTPKEPAPTTPKKPA--PT
                                                                                                                                                                                                                                                                                                                                                                        ELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA- 591
                                                                                                                            KETATTTEKTTESKITATTT
                                                                                                                                                     PSYP--PTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAKPSYPPTHKAKPT
                                                                                                                                                                                    PKEPGVPTTKTPAA-----TKPEMTTTAKDKTTERDLRTTPETTTAAP------KMT
                                                                                                                                                                                                                                                                                   PPTYKAKPTYK-AKPTYKAKPTYPSTYKAKPSYSPTYKAKPSYP----PTYKAKPSYPPT
                                                                                                                                                                                                                                                                                                                                                   SYPPTYKAKPTYPSTYKAKPTYK--ARPSYP----PTYKAKPS-----YPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                                  SYPPTYKPKISYPPTYKAKPSYPSTYKAKSSYPPTYKAKPSYPSTYKAKSSYPPTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEPAPTTPKEPAPTT----PKEPAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKAKPSYPPTYKAKPTYPSTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKVKPTYKAKPTYPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YPSTYKAKPSYT---YPSTYKAKPTYP----PTYKAKPSYPPTYKAKPTYKAKPTYKAKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -PTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPA---PTTPKEPA----PTTPKKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTPTTIKSAPT-TPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA--P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -APTTPKEPA--PTAPKKPA--PTTPKEPA--PTTPKEPA--PTTTKEPS--PTTPKEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTYKPKPSYP--PSYKTKKTYPPTYKPKLTYPPTYKPKPSYPPSYKPKKTYPPTYKPKLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SYPPTYKSKPTYKPKTTYPPTYKAKPSYPPTYKPKKTYPPT-YKPKLTYPPTYKPKASYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----YPPTYK--AKPS-----YPPTYKPKKTYPPTYKPKLTYPPTYKPKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKE----PASTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTTKSAPTTPKEPAPTTPKKPAPT-TPKEPAPTT----PKEPTPTTPKEPAPTTKEP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     al Similarity
239; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10.1%; Score 506.5; DB 6; 29.9%; Pred. No. 7.8e-25; tive 78; Mismatches 308;
                                                                                                                              759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                     601
                                                                                                                                                                                                                                                                                                                                                                                                                    498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53
```

밁 Qy D δ

089

δô В

밁 Ş 밁 S

> 494 444 434 418

õ 밁

밁 õ В õ 멍 ő ₽ õ 밁 Ş В Š 밁 Ş 밁 Ş

Run on: OM protein

Maximum Minimum

DB DB

```
Title:
Perfect score:
Séquence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
417
417
413.5
413.5
404.
                                                                                                                                                                                                                                                                                                                                                                                                                                   488.5
488.5
476.5
466.5
463.5
                                                                                                                                                                                                                                                                                          423.5
423.5
423.5
423.5
423.5
423.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             506.5
498.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Issued_Patents_AA:*
1: /cgn2_6/ptodata/2,
2: /cgn2_6/ptodata/2,
3: /cgn2_6/ptodata/2/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-556-246-1_COPY_200_1140
5011 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           April 26, 2002, 16:13:29; Search time 49.78 Seconds (without alignments) 425.384 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gapop 10.0 ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VKDNKKNRTKKKPTPKPPVV......DMDYLPRVPNQGIIINPMLS 941
                                                                                                                       /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/B_COMB.pep:*
/cgn2_6/ptodata/2/iaa/Backfiles1.pep:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000000000
                                                                                                                                                                                                                                             1837
826
826
652
1721
1721
1721
1767
1867
2035
2035
2476
1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gapext 0.5
                        US-08-700-651-5
US-08-928-3618-6
US-09-219-849-5
US-09-219-849-5
US-09-083-116-5
US-09-083-116-5
US-09-083-116-2
US-09-083-116-2
US-09-041-886-23
US-09-041-886-23
US-08-479-533A-53
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-477-598-103
US-08-397-633A-53
US-08-397-633A-53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-103-429A-3
US-08-928-361B-5
US-07-638-431-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-103-429A-4
5202236-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5202236-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US92-00018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                           July 2236

July 2236

July 2236

July 2236

July 2336

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    212252
                                                                                                   Sequence 23, Appl
Sequence 132, App
Sequence 53, Appl
Sequence 68, Appl
Sequence 103, App
Sequence 101, App
Sequence 75, Appl
Sequence 75, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                                                                      Sequence
Sequence
Sequence
Sequence
Sequence
              Sequence
Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 4,
                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                               5, Appli
6, Appli
5, Appli
5, Appli
5, Appli
5, Appli
2, Appli
2, Appli
                 50, Appl
2, Appli
19, Appl
65, Appl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4, Appli
5202236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 1
US-09-103-429A-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 4, Application US/09103429A
Patent No. 6187558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENERAL INFORMATION:
```

Result

10 111 112 13 14 15 16 116 117 118 119 20 21 22 23 23 24 25 26

4.5	44	43	4 2	٠ ٠	41	40	39	38	37	30	וו	S)	34	ω ω	32		21	30	2.9	0 6	ာ
367.5	367.5	367.5	1 0	20 1	368	369.5	370.5	371	371	0/1	0 1	374.5	376	381	185		7 80 E	398.5	404.5	1	404 5
7.3	7.3	1.3		7	7.3	7.4	7.4	7.4	7.4		7 1	7.5	7.5	7.6	1.0	1 0	ж Э	8.0	0.1	9 !	8 . 1
/811	1187	1187	1 1 0	762	762	1231	1064	762	162	1 -	761	1537	960	1848	1840	2 1	682	682		100	408
U	ı	-	4	_	μ,	w	_	٢	-		J	Н	4	U	1 #	-	Н	ш	4	_	4
PCT-0894-04490 20	US-U8-040-0600-20		115-09-330-559-38	US-08-397-633A-31	US-08-642-255-120	US-08-904-263A-4	US-08-642-255-62	US-08-39/-633A-26	05-00-044-000 F14	00 00 CAD-DEE-11/	IIS-08-707-237A-84	US-08-325-267A-2	US-09-219-849-6	PCT-0895-10601A-0	00 00 600 704 0	116-08-306-701-6	US-08-397-633A-36	US-08-642-255-126	00 00 00 00 00 00 00 00 00 00 00 00 00	TIS-08-478-029A-65	US-08-475-411A-65
ļ	-	28	Sequence 28, Appl		1 1 0	Sequence 4, APPT	Sequence oz, appr			Seguence 114, App	Sequence 84, APPI		ى ر			Sequence 6, Appl1	sequence so, Appr			Sequence 65, Appl	Sequence 65, Appl

ALIGNMENTS

Granados, Robert R

```
; ORGANISM: Tr
; TISSUE TYPE:
US-09-103-429A-4
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                 TELEFAX: (607) 256-36
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                          APPLICATION NUMBER: US/09/103,429A
FILING DATE: 24-JUN-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A
REGISTRATION NUMBER: 34,390
REGISTRATION NUMBER: 34,390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Granados, Robert R
APPLICANT: Wang, Ping
TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin
TITLE OF INVENTION: CDNA and Related Products and Methods
                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: BT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (607) 256-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUMBER OF SEQUENCES:
                                                                                                                                                        MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-SENSE: NO
                                                                                                                                      ORIGINAL SOURCE:
                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 118 I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                   TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: Brown, Pinnisi & Michaels, P.C.
118 No. 6187558th Tioga
                                                                                                                                                                                                                                                                            805 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                    linear
                                                                                               Trichoplusia ni
E: peritrophic membrane
                                                                                                                                                                                                                                       single
                                                                                                                                                                                                                                                                                                                                      256-3628
    10.4%; 27.8%;
                                                                                                                                                                                                                                                                                                                                                                                                    BTI-39
        Score 522; DB 4;
Pred. No. 8.9e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Version
                          Length 805;
```

	y i		
		Service Control of the Control of th	
	\$		
	*		
	. 8		

	3.		
	The state of the s		
4			
4	11 11 11 11 11 11 11 11 11 11 11 11 11		
		THE ACCOUNT ASSESSMENT OF THE PROPERTY OF THE	

Search completed: April 26, 2002, 16:19:03 Job time: 328 sec

```
RESULT
NFH_RAT
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                  This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Enthe European Bioinformatics Institute. There are no restricted the European Bioinformatics institute.
                                                                                                                                                                                                                                                       +
                                                                                                                                                                                                                                                                                                                                          ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-89184647; PubMed-2028342;
Lieberburg I., Spinner N., Snyder S., Anderson J., Goldgaber D.,
Smulowitz M., Carroll Z., Emanuel B.S., Breitner J., Rubin L.;
"Cloning of a cDNA encoding the rat high molecular weight
neurofilament peptide (NF-H): developmental and tissue expression
the rat, and mapping of its human homologue to chromosomes 1 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-89065087; PubMed-3143606; Breen K.C., Robinson P.A., Wion D., Anderton B.H.; Partial sequence of the rat heavy neurofilament polypeptide (NF-H) Identification of putative phosphorylation sites."; PEBS Lett. 241:213-218(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "The lar
in situ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RObinson P.A., Wion D., Anderton B.H.; "Isolation of a cDNA for the rat heavy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-89 AND 243-313
MEDLINE-87080760; PubMed=2878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dautigny A
Jolles P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P16884; Q63368;
01-AUG-1990 (Rel. 15, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
REUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT (NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H) (FRAGMENT).
NEFH OR NFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 318-831 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEBS Lett. 209:203-205(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 37-831 FROM N.A. MEDLINE-88309090; PubMed-245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (NF-H)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rattus norvegicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NFH_RAT
                                                                                                                                  INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTE OF AXONAL CALLBER.

OF AXONAL CALLBER.

OF AXONAL CALLBER.

OF AYONAL CALLBER.

THE PROST OF THE FUNCTION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTION THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.

SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CAUTION: REF. 1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITIC ONWARD AND IS LONGER DUE TO A FRAMESHIFT.
                                                                                                                                                                                                                                                                           FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, MAND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT SUBSERVED BY THE TWO SMALLER NF PROTEINS.

PTM: THERE ARE A NUMBER OF REPEARS OF THE TRIPEPTIDE K-S-P, NFH PHOSEHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   large neurofilament subunit (NF-H)
                                                                                                                                                                                                                                                                                                                                                                                                                      Natl. Acad. Sci. U.S.A. 86:2463-2467(1989)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          detection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Biophys.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A., Pham-Dinh D., Roussel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PubMed=2878828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Res. Commun.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PubMed=2457365;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chordata;
Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Craniata; Vertebrata; I
Sciurognathi; Muridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  154:1099-1106(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ი,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             neurofilament polypeptide
                                                    There are no rest
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Felix J.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    rat:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Nussbaum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cDNA cloning and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Euteleostomi
                                                                      restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Murinae;
                                                                                                                                                                                                                              NF-H), THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN)
                                                                                          EMBL
                                                                                                                                                           POSITION
                                                                                                                                                                                                                                                                               MAINTENANCE
                                                                                                          þ
                                                                                      collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     J.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'n
                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                  z
                                                                                                                                                                                                                                                                                                                                 SI
                                                                      its
```

```
Ş
                        B
                                        δÃ
                                                             맑
                                                                              Q
                                                                                                    Вþ
                                                                                                                     ρy
                                                                                                                                          В
                                                                                                                                                            δÃ
                                                                                                                                                                                В
                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                         Ωy
                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                             20
                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M37227; AAA41693.1; & EMBL; X13804; CAA32038.1; & EMBL; M21964; AAA41695.1; & EMBL; J04517; AAA41692.1; - EMBL; J04517; AAA41692.1; - PIR; A30796; A30796.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR;
    510
                        645
                                                             589
                                                                                 390
                                                                                                    529
                                                                                                                                           471
                                                                                                                                                                                 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
                                                                                                                        339
                                                                                                                                                               288
                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Intermediate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Piam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001
                                                                                                                                                                                                     246
                                                                                                                                                                                                                        353
                                                                                                                                                                                                                                           204
                                                                                                                                                                                                                                                               294
                                                                                                                                                                                                                                                                                   153
                                                                                                                                                                                                                                                                                                                                          197 PSMST--HIKVKSEEKIKVVE----
                                                                                                                                                                                                                                                                                                                          97
                                                                                                                                                                                                                                                                                                                                                          37 PDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTNKQ
                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
mes 195; Conserv
                                                        A25649;
B25649;
S02003;
PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPT
                  ---KAKEPPKKVEEEKTPATPKTEVKESKKDEAPKEAQKP-KAEEKEPLTEKPKDSPGEA
                                   APTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKETAPTT
                                                                                               VKEGAKSLAEAKSPEKAKSPVKEEIKPPAEVKSPEKAKSPMRKEAKSPEKAKTLDVKSPE
                                                                                                                                      AKSPAEA--KSPAGAKPPAEAKSPAEAKSPAEAKSPAEAKSPAEAKSPVEVKSPEKAKSP
                                                                                                                                                                                             --PTTPKEPA----PTTPKEP----PTA
                                                                                                                  APTITKS----APTIPKEPSPTTTKEPA----PTTPKEPAPTTPKKPAPTTPKEPAPTT
                                                                                                                                                        PKKPAPTTPKEPAPTTTKEPSPTTPKEPA----PTTTKS-----APTTTKEP
                                                                                                                                                                            KSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPA-SVKSPSEAKSPAGAKSPAEAKSPVV
                                                                                                                                                                                                                   SPGEAKSPAEAKSPAEVKSPAEAKSPAEAKSPASVKS-PGEAKSPAEAKSPAEVKSPATV
                                                                                                                                                                                                                                EP----APTTPKEPA----PTTTKEPA----PTTTKSAPTTPKEPA----PTTPKKPA---
                                                                                                                                                                                                                                                          PAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKS-PAEAKSPAEVKSPATVK
                                                                                                                                                                                                                                                                            PTTPKEPA-STTPKEPTPTTIKSAPTTPKEPA----PTTTKSAPTTPKEPA----PTTTK
                                                                                                                                                                                                                                                                                                KEAQGEEEEEAEEGGEEAATTS-----PPAEEAASP----EKETKSPVKEEAKSPAEAKS
                                                                                                                                                                                                                                                                                                                   TSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF00038; filament; 1.
TE; PS00226; IF; 1.
mediate filament; Coiled
                                                                                                                                                                                                                                                                                                                                                                                                                                   831 AA;
                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                               276
164
185
193
193
346
373
373
482
482
757
757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1664; IF
                                                                                                                                                                                                                                                                                                                                                                                           9.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                             641
164
185
193
193
199
199
346
373
482
482
7591
7591
7759
                                                                                                                                                                                                                                                                                                                                                                                                                                   89486
                                                                                                                                                                                                                                                                                                                                                                                                                                   MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALT_FRAME.
                                                                                                                                                                                                                                                                                                                                                                                59;
                                                                                                                                                                                                                                                                                                                                                                                                                                       L -> T (IN REF. 2)

M -> Y (IN REF. 1)

K -> N (IN REF. 1)

A -> V (IN REF. 1)

A -> V (IN REF. 2)

P -> S (IN REF. 2)

RK -> KE (IN REF. 2)

RK -> KE (IN REF. 2)

A -> V (IN REF. 2)

A -> V (IN REF. 2)

A -> D (IN REF. 2)

R -> D (IN REF. 2)

R -> D (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                           Score 497;
Pred. No. 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             51 x
I ->
                                                                                                                                                                                                                                                                                                                                        -----KSEKETVIVEEQTEEIQVTEEVTEEED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     coil;
                                                                                                                                                                                                                                                                                                                                                                                                                                 1B0973C3F13EF768
                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X 3 AA TANDEM RI
I (IN REF. 2)
S (IN REF. 2)
S (IN REF. 2)
T (IN REF. 2)
Y (IN REF. 1)
V (IN REF. 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Neurone; Phosphorylation; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                           REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          REF. 4).
(IN REF. 4).
                                                                                                                                                                                                                                                                                                                                                                                         .3e-13;
                                                                                                                                                                                                                                                                                                                                                                                                    DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      . 2)
                                                                                                                                                                                                                                                                                                                                                                                307;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEATS OF K-S-P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AND
                                                                                                                                                                                                                                                                                                                                                                                                                               AND 4).
8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AND
                                                                                                                                                                                                                                                                                                                                                                                                 Length 831;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AND 4).
                                                                                                                                                                                                                                                                                                                                                                                Indels 146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           $$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4).
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                   700
                                                                             449
                                                         644
                                                                                                588
                                                                                                                   389
                                                                                                                                       528
                                                                                                                                                         338
                                                                                                                                                                            470
                                                                                                                                                                                                                   411
                                                                                                                                                                                                                                                           352
                                                                                                                                                                                                                                                                                                                                       242
                                                                                                                                                                                                287
                                                                                                                                                                                                                                      245
                                                                                                                                                                                                                                                                             203
                                                                                                                                                                                                                                                                                                293
                                                                                                                                                                                                                                                                                                                    152
                                                                                                                                                                                                                                                                                                                                                            96
                                                                                                                                                                                                                                                                                                                                                                              35;
```

```
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Вb
RESULT
RPB1_M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RPB1_MOUSE
P08775;
01-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Corden J.L., Cadena D.L., Ahearn J.M. Jr., Dahmus "A unique structure at the carbox1 terminus of t of eukaryotic RNA polymerase II."; proc. Natl. Acad. Sci. U.S.A. 82:7934-7938(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIINE-87280135; PubMed=3038894; Ahearn J.M. Jr., Bartolomei M.S., West M.L., "Cloning and sequence analysis of the mouse the largest subunit of RNA polymerase II."; J. Biol. Chem. 262:10695-10705(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus musculus (Mouse)
Eukaryota; Metazoa; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAR-1992 (Rel. 21, Last sequence update) 20-AUG-2001 (Rel. 40, Last annotation update) DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT POLR2A OR RPO2-1 OR RPII215.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1829
                                                                                                                                                                                                                                                                                                                                                                                           Wintzerith M., Acker J., Vicaire S., Vigneron M., Kedinger C.; "Complete sequence of the human RNA polymerase II largest subunit."; Nucleic Acids Res. 20:910-910(1992).

-i- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=86068017; PubMed=2999785;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1587-1970 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=92178992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REVISIONS, AND PRESENCE OF AN ADDITIONAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               383
                                 PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED. THE PHOSPHORYLATION ACTIVATES POLZ.

MISCELLANGOUS: THREE DISTINCT ZING-CONTAINING RNA POLYMERASES ARE FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
                                                                                                                                                                                                                                                                                             CATALYTIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SDEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SY-SPTSPSY-SPTSPSY-SPTSPNY-SPTSPNYT-PTSP-SYSPTSP-SYSPTSPNY-T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPA : | | : | | : | : | : | : | : | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
                                                                                                                                                                                SUBCELLULAR LOCATION: NUCLEAR
                                                                                                                                                                                                                  SUBUNIT: RNA POLYMERASE
                                                                                                                                                                                                                                                                                                                                                                 OF DNA INTO RNA USING THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAPTTPKETAPTTPKEP--APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTSPTY-SPTTPKY-SPTSP--TYSPTSPVYTPTSPK-----YSPTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT-STTSDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SP-SYSPTSPKYT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TPKETAPTTPKGTAPTTLKEPAPTTPK-KPAPKELAPTTTKEPTSTTSDKPAPTTPKGTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTSP-NYSPTSPSY-SPTSPSY-SPTSP-SYSPSSPR----YTPQSPT-YTPSSPSYSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEP--APT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -SPTSPSYSPTSPSY-SPTSPSYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSP-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPS-SPTSPS-SPTSPS-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               YSPTSP-TYSPTSPKGSTYSPTSPGY-SPTSP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Rel.
                                                                                                                                                                                                                                                                                             ACTIVITY: N NUCLEOSIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
           POLYMERASE II FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PubMed=1542581;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chordata;
Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         II CONSISTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Craniata; Vertebrata; Sciurognathi; Muridae;
                                                                                                                                                                                                                                                                                                                                                                     FOUR
       POLYMERASE I FOR THE THE MRNA PRECURSOR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     West M.L.,
                                                                                                                                                                                                                                                                                                                                                                     RIBONUCLEOSIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1970
                                                                                                                                                                                                                                                                                             TRIPHOSPHATE =
                                                                                                                                                                                                                          OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EXON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Dahmus M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -TYSLTSPAI-
                                                                                                                                                                                                                             12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genomic locus encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (EC
                                                                                                                                                                                                                                                                                                                                                                 TRIPHOSPHATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2.7.7.6)
                                                                                                                                                                                                                                                                                             z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .
H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   largest subunit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Murinae;
           AND
                                                                                                                                                                                                                                                                                             PYROPHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTYSPTSPK
                                                                                                                                                                                                                             SUBUNITS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RPB1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SPDD
                                                                                                                                                                                                                                                                                                                                                                         AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              676
```

```
SOFTITION AND CONTRACTOR OF THE WAY AND CONT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δδ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                    20
                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
Qy
                                                   В
                                                                                        QУ
                                                                                                                                           멍
                                                                                                                                                                                    QΥ
                                                                                                                                                                                                                                     рь
                                                                                                                                                                                                                                                                               ΔÃ
                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWI
between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00115; RNA_POL_II_REPEAT; 42.
Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.
ZN_FING 71 87 C2H2-TYPE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MGD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000684; RNA_polII_repeat.
InterPro; IPR000722; RNA_pol_A.
InterPro; IPR002879; RNA_pol_A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01854;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                        1829
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1625
                                                                                                                                                1876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  151
                                                        1923
                                                                                                                                                                                                                                                                                                                                                                                 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                263
         677
                                                                                                                                                                                            560
                                                                                                                                                                                                                                                                                      501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          III FOR 5S AND TRNA SIMILARITY: BELONGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MGI:98086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SWISS-PROT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPTTPKEPASTTPKE-PTPTTIKSAPTTPKEPAPTTTKSAPTTPKE----PAPTTTKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A28490; A28490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     M12130;
M14101;
                                                                                                                                                                                                                                                                                                                                                                                                                             SY-SPTSPSY-SPTSPSY-SPTSPNY-SPTSPNYT-PTSP-SYSPTSP-SYSPTSPNY-T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTPGSPGSPGPSSPYIPSPGGAMSPSYSPTSPA-YEPRSPGGYTPQSPSYSPTSPSY-SP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEP--APTTPKEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PSPMGGISPAMTPWNQGATPAYGAWSPSVGSGMTPGAAGFSPSAASDASGFSPGYSPAWS
                                                                                                                                                                                                                                                                                 TPKETAPTTPKGTAPTTLKEPAPTTPK-KPAPKELAPTTTKEPTSTTSDKPAPTTPKGTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TSPSY-SPTSPNYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-
         STPE
                                                                                            PAPTTPKETAPTTPKEP--APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDE
                                                                                                                                                                                       PTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT
                                                                                                                                                                                                                                                                                                                                  PTSP-NYSPTSPSY-SPTSPSY-SPTSP-SYSPSSPR----YTPQSPT-YTPSSPSYSPS
                                                                                                                                                                                                                                                                                                                                                                PTTPEELAPTTPEEPTTTPEEPAPTTPKAAAAPNTPKEPAPTTPKEP--APT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPA
                                                     YSPTSP-TYSPTSPKGSTYSPTSPGY-SPTSP
                                                                                                                                                PTSPTY-SPTTPKY-SPTSP--TYSPTSPVYTPTSPK
                                                                                                                                                                                                                                          SP-SYSPTSPKYT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -SPTSPSYSPTSPSY-SPTSPSYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 180;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF00623; RNA_pol_A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
         680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1499
1970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA40071.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA40071.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RNA_pol_A2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        entry is copyright. It is produced through iss Institute of Bioinformatics and the EM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1536
AA; 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1958
1498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9.9%;
33.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 217175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENES.
TO THE RNA POLYMERASE BETA' CHAIN FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101;
                                                                                                                                                                                                                                   - PTSPSYSPSSPEYTPASPK - YSPTSPKYSPTSPK - YS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 498.5;
Pred. No. 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P -> R (IN REF. 1 AND 2).
MISSING (IN REF. 1 AND 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOXYL-TERMINAL 7-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7D76F38FD92A657E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.9e-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         , 1;
13;
                                                                                                                                                     --YSPTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESIDUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                       PTYSPTSPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -SPTSPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95;
                                                                                                                                                                                              STTSDK
                                                             -SPDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                   618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204
                                                                                                                                                                                                                                                    1875
                                                                                                                                                                                                                                                                                                                                              1828
                                                                                                                                                                                                                                                                                                                                                                                                                                         1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1566
                                                                                                           676
                                                                                                                                                          1922
                                                                                                                                                                                                                                                                                                559
                                                                                                                                                                                                                                                                                                                                                                                             500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    262
                                                               1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47;
```

멍

1966

SDEE 1969

```
RESULT
RPB1_HU
ID RP
AC P2
DT 01
DT 01
DT 01
DT 00
DF DN
GN PC
OC EU
OC Ma
OC 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                       멼
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                            RPB1_HUMAN
P24928;
01-MAR-1992
01-MAR-1992
01-MAR-1992
20-AUG-2001
                                                                                                                                                                                                                                                                                                                                       HUMAN
SEQUENCE FROM N.A. MEDLINE-92178992;
                                                                                                                      Eukaryota;
                                                                                                                                                                      POLR2A.
                                                                                             Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                          DNA-DIRECTED
                                                                         NCBI_TaxID-9606
                                                                                                                                             Homo sapiens (Human
                                                                                                                                                                                                                                                                                                                                                                                                                                       567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTTTKATTTTSGECKME----PSKRADCGYPGITESQCRSKGCCFDSSIPQTKWCFYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MPRVRKPKTTPTPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KVTTTKKTITTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPK----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAP
                                                                                                                                                                                                                                                                                                                                                                                                                                  LPQVADCKVAPSSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -TTTPTTTTTP-----TTTTTKAT------TTTPTTTTTTTTTKATTTTPTTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTTTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ISGTKWCFYSTSQVAA - - TKTTTTPTTTTTTTTTTTTKATTTTP - - - - - - TTTTTTPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TKGCCFDSSIPQTKWCFYTLSQVADCKVEPSQRVDCGFRGIT----ADQCRQKNCCFDSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PKG----TAPTT-----PKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VKPSKREMCGSKGITKKQ----CKKKNCCFDPKGHGGIHCFHRKPKGHSHEEHTTTTK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GKAPATAAAPVPTTAASKAPTTAAAATHSTAAAAAPTTAASAAKSKERSTSSSSEEEHCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TINKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPAPTTTKEPAPTTTKS-APTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --PTTKEPAPT---TPKEPAPTAPKKPAPTTPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------AAATAAAETTAAAGEAPTTTAPATTAAGKAPTTAAATAPTTAAAGAPTTAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTAAAVAATGKDTTAAAEGSAAAEKTAA----AGEVSAPPT--AAVAATGEDATT----
                                                                                                                                                                                   (Rel. 21, Created)
(Rel. 21, Last sequence update)
(Rel. 40, Last annotation update)
D RNA POLYMERASE II LARGEST SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -TTTTTKATTTPTTTTTTP---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -APTTIQIATTTT - - - TPTTT
                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -TTTTTKATTTTTTSGECKMEPSK-----REDCGYSGITESQCR
PubMed=1542581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -TTTTKATPTTT-------TTTKATTTTT----TPTT----
                                                                                          Chordata;
Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                  580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             854
                                                                                        Craniata; Vertebrata; Catarrhini; Hominidae;
                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                         1970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTTTKATPTT - - - - - - -
                                                                                                                                                                                     OEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --EPAPTTPKEPAPTTTKE
                                                                                                                                                                                2.7.7.6) (RPB1).
                                                                                        Euteleostomi;
Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   151
```

```
Q
                                                                         В
                                                                                                                             Qy
                                                                                                                                                                                                                                             δÃ
                                                                                                                                                                                             B
                                                                                                                                                                                                                                                                                                          Вb
                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOFTE TWO DRAWERS BY BURNESS OF THE TWO DRAWS BY BURNESS OF THE TWO B
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
Matches 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR0028/5, ...

InterPro; IPR0028/5, ...

InterPro; IPR0028/5, ...

Pfam; PF00823; RNA_pol_A; 1.

Pfam; PF01854; RNA_pol_A2; 1.

Pfam; PF01854; RNA_POL_II_REPEAT; 43.

Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Transferase; DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.

DNA-binding; Nuclear protein; Phosphorylation; Zinc-finger.

C2H2-TYPE (POTENTIAL).

CARBOXYL-TERMINAL 7-RESIDUE REPEA

CARBOXYL-TERMINAL 7-RESIDUE REPEA
                                                                         1625
                                                                                                                                                                                                                                                                                                       1507
                                                                                                                                                                                             1567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000684; RNA_polII_repeat.
InterPro; IPR000722; RNA_pol_A.
InterPro; IPR002879; RNA_pol_A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; S21054; S21054.
MIM; 180660; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-95347616; PubMed-7622068;
Mita K., Tsuji H., Morimyo M., Takahashi E., Nenoi M.,
Ichimura S., Yamauchi M., Hongo E., Hayashi A.;
"The human gene encoding the largest subunit of RNA polymerase II.";
Gene 159:285-286(1995).
-i- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTIC
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
                  323
                                                                                                                                                                                                                                                 205
                                                                                                                                                                                                                                                                                                                                                                 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wintzerith M., Acker J., Vicaire S., Vigneron M., Kedinger C.; "Complete sequence of the human RNA polymerase II largest subunit."; Nucleic Acids Res. 20:910-910(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBCELLULAR LOCATION: NUCLEAR.

PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
THE PHOSPHORYLATION ACTIVATES POLZ.
MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES AF
FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
III FOR 5S AND TRNA GENES.
SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          x63564; CAA45125.1;
x74874; CAA52862.1;
x74873; CAA52862.1;
x74872; CAA52862.1;
x74871; CAA52862.1;
x74870; CAA52862.1;
x74870; CAA52862.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE
PAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTP
                                                                                                                                                                                                                                                                                                                                            PTPTTPKEPASTTPKE-PTPTTIKSAPTTPKEPAPTTTKSAPTTPKE-----PAPTTTKE 204
                                                                TSPSY-SPTSPNYSPTSPSY-SPTSPSY-SPTSPSY----
                                                                                                    TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTTKEPSPTTPKE 322
                                                                                                                                                                       PTPGSPGSPGPSSPYIPSPGGAMSPSYSPTSPA-YEPRSPGGYTPQSPSYSPTSPSY-SP 1624
                                                                                                                                                                                                                           PAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEP--APTTPKEPTP 262
                                                                                                                                                                                                                                                                                     PSPMGGISPAMTPWNQGATPAYGAWSPSVGSGMTPGAAGFSPSAASDASGFSPGYSPAWS 1566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: RNA POLYMERASE II CONSISTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RNA(N)
                                                                                                                                                                                                                                                                                                                                                                                                                     181;
                                                                                                                                                                                                                                                                                                                                                                                                               Similarity 33...
81; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1958 CARBOXYL-TERMINAL 7-F
1067 W -> L (IN REF. 2).
1449 D -> Y (IN REF. 2).
A; 217205 MW; 6876FC25692A657E
                                                                                                                                                                                                                                                                                                                                                                                                                                              10.0%;
                                                                                                                                                                                                                                                                                                                                                                                                               101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                         Score 503.5; I
Pred. No. 2.5e
01; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                           .5e-13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                            167;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DIFFERENT
                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBUNITS
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEATS
                                                             SPTSPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                         95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                             1671
                                                                                                                                                                                                                                                                                                                                                                                                      47;
```

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            뮰
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DЪ
δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MUC1_XENLA
Q05049;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-1994 (Rel. 30, Created)
01-OCT-1994 (Rel. 30, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
01-OCT-1994 (Rel. 30, Last annotation update)
INTEGUMENTARY MUCIN C.1 (FIM-C.1) (FRACMENT).
Xenopus laevis (African clawed frog).
Xenopus laevis (African clawed frog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       XENLA
                                                                                                                                                                      This
EMBL; L02115; AAA74725.1;
PIR; A45155; A45155.
HSSP; P04002; 1WFA.
                                                                                                                                                                                                                                                                                                                                            polymorphism.";
J. Biol. Chem.
                                                                                                                                                                                                                                                                                                                                                                            Hauser F., Hoffmann W.;
"P-domains as shuffled cysteine-rich modules in integumentary C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 470
                                                                                                                                                                                                                                                                                            -
                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=93077556; PubMed=1447205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  rissue-skin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=8355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Xenopodinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           736
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         786
                                                                                                                                                                                                      BIOL. Chem. 267:24620-24624(1992).

BIOL. Chem. 267:24620-24624(1992).

BUNCTIONS: COULD BE INVOLVED IN DEFENSE AGAINST ME PUNCTIONS: PROTECTS THE EPITHELIA FROM EXTERNAL INFECTIONS. SECRETED.

SUBCELLULAR LOCATION: SECRETED.

ALTERNATIVE PRODUCTS: A NUMBER OF DIFFERENT FORM MAY BE PRODUCED BY ALTERNATIVE SPLICING.

TISSUE SPECIFICITY: SKIN.

PTM: EXTENSIVELY O-GLYCOSYLATED.

PTM: EXTENSIVELY O-GLYCOSYLATED.

SIMILARITY: CONTAINS 6 P-TYPE (TREFOIL) DOMAINS.
                                                                                                                                  s SWISS-PROT entry is copyright. It is produced through a collaboration ween the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Y---PPSYKPKKTYPPTYK-----PKISYPPTYK----TKPSYPASYKRKTSYPPTYKPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTT--PKETAPTTPKEPAPTTPK--KPAPTTPET-----PPPTTSEVSTPTTTKEPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTYKPKITYPPTYK-----PKPSYPPSYKPKTTYPPTYKPKIRYPPTYKPKASYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TT--PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPT-TPKGTAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTYKPKPSYPPTYKPKTTYPPTYKRKPSYPTPYKQKPSYPPTYKSKSSYPTSYKSKKTYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MPRVRKPKTT---PTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ---KPTKAPKKP-TSTKKPK-T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKT----TERDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PPTYKPKITYPPTYK-PKPSYPPSYRPKITYPPTYKPKKSYPQAYKSKGSYPPSYQPKKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP----KKPAPKELAPTTTKGPTSTTSDKPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YPPTYKPKISYPPTYKPKITYPPTYKPKISYPPAYKPKISYPSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ISYPSTYKAKP----
                                                                         non-profit institutions as long as its content is in and this statement is not removed. Usage by and for company and the statement is not removed. Usage by and for company and email to agreement (see the
                                                                   equires a license agreement (S email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Xenopus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----RYPPTYKPKP-----SYASSYKPKIRYPPTYKPKPSYASSYKPKIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -PKITYPPTYK-PKPSYPTPYKQKPSYPPIYKSKSSYPTAYKSKKTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -SYPPTYKPKPSYASSYKPKIRYPPTYKPKPSYASSYK----PK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          662
                                                                                         (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A
                                                                                                                                                                                                                                                                                                               AGAINST MICROBIAL
M EXTERNAL ENVIRON
                                                                                                                                                                                                                                                                                 FORMS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pipidae;
                                                                                                                                                                                                                                                                                                                 ENVIRONMENT
                                                                                                                                                                                                                                                                                 THE PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                       mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                827
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  785
```

```
Repeat; NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR000519; P_trefoil.
pfam; PF00088; trefoil; 6.
SMART; SM00018; P; 6.
                                                                                     DOMAIN
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE;
                                                                                                                                                                                                                                                         REPEAT
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                               DISULFID VARIANT
                                                       DISULFID DISULFID DISULFID
                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amphibian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PS00025;
                                                                                                                                                                                                                          137
161
218
218
225
240
250
260
276
276
276
276
276
306
353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81
89
97
105
113
121
129
                                                                                                                                                                                                                                                                                                                                                                       402
412
420
432
skin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P_TREFOIL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            144
                                                                                                                                                                                                                                                                                                                                                                                                                             Glycoprotein; Alternative
                                                                                                                                                                                                                                                                                                                                                                                                                 1-8.
1-8.

P-TYPE 1.

8 X APPROXIMATE TANDEM REPEATS, THR-RICH.
2-1.
2-1.
2-2.
2-3.
2-4.
2-5.
2-6.
2-6.
2-7.
2-8.
2-7.
2-8.
2-7.
2-8.
2-7.
2-8.
2-12 X APPROXIMATE TANDEM REPEATS,
THR-RICH.
2-12.
THR-RICH.
    MW;
              3-1.
3-2.
3-3.
3-3.
3-4.
3-6.
3-6.
3-6.
3-7.
3-7.
3-10.
3-110.
3-110.
3-110.
3-12.
3-12.
3-12.
3-12.
3-12.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-110.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3-12.
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6
                                            Y SIMILARITY.
   F085277F1ED2FD40 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA APPROXIMATE TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   splicing
```

Query Match Best Local Sim Matches 223;

Similarity

10.2%; 28.1%;

Score 509.5; DB 1; Pred. No. 6.3e-14;

Length Indels

662;

Gaps

30;

Conservative

48;

Mismatches

```
RESULT 11
FP1_WTCO
ID FP1_W
AC 02543
DT 30-MA
DT 30-MA
DT 30-MA
DT 30-MB
EP1.
OS MYtill
OC EUKar
OC MYtill
OX NCBI_RN (11)
RP SEQUE
RX MEDIJ
RA ITOUE
RA MEDIJ
RA ITOUE
RA MEDIJ
RA ITOUE
RA MEDIJ
RA ITOUE
RA MEDIJ
RA ITOUE
                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           망
                                                                                                                   Eukaryota; Metazoa; Mol
Mytiloidea; Mytilidae;
NCBI_TaxID=42192;
                                                                                                                                                                                                                                                           FP1_MYTCO
Q25434;
                                                                                                                                         Mytilus coruscus (Sea mussel).
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia;
                                 Inoue K., Ta
Harayama S.;
                                          MEDLINE-96394686; PubMed-8798340; Inoue K., Takeuchi Y., Takeyama S.,
                                                                                                                                                                                          30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
ADHESIVE PLAQUE MATRIX PROTEIN PRECURSOR (FOOT
                                                                                TISSUE-Foot;
                                                                                          SEQUENCE FROM N.A.
Adhesive protein cDNA sequence of the ts evolutionary implications.";
                                                                                                                                                                                                                                                                                                                                                   814
                                                                                                                                                                                                                                                                                                                                                                                  853
                                                                                                                                                                                                                                                                                                                                                                                                              758
                                                                                                                                                                                                                                                                                                                                                                                                                                            793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               280
                                                                                                                                                                                                                                                                                                                                            LRDLQTTDVSLLAIAATLDAIGEKLKDQKARNQQVMDRLCEIEKILGPPKS
                                                                                                                                                                                                                                                                                                                                                                      PRKMTSTMPELNPTSRIAEAM----LQTTTRPNQTPNSKLVEVN-----PKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                               DIAIPVIDPPVPQEIAVAEIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAA
                                                                                                                                                                                                                                                                                                                                                                                                    EVPITAGDNPDNTSVGISEVVPTIAEKPVEEVPTS--EIPEQSSSPS--DSVPVAKITPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TVE-FPEAVAEKVLDPAI--TEAPVTTQEPDVANINDGAPATEI-TTPAVEIVTAAAEVS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTP--ETTTAAPKMT
                                                                                                                                                                                                                                                                                                                                                                                                                                 ITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KETATTTEKTTESKITATTTQVTST-----TTQDTTPFKITTLKTTTLA-PKVTTTKKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EVAVAPITAPEPIP----EPEPSLATPTEPIPVEAPVVIQEAVDAVEVPVTETSTSIPET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPEAAADLIIEPVEPPAPIPDLLEQTTSVPAVEAAESTSS--PIPETSLPPPNEAVASP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TLKEPA-----PTTPKKPAPKELAPTTT----KGPTSTTSDKPAPTT----PKETAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LAPTTTKEPTSTTSDKPAPTTPKGTAPTTP-----KEPAPTTPKEP--APTTPKGTAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               V----LPPVAAEPVPAVVAEETPETPAPASAPVTIAALDIPEVAPVIAAPSDAPAEAPSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TPKEPAPTTPKEPAPTTPKETAPTTP-KGTAPTTLKE------PAPTTPKKPAPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPD-----ESTPEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAPIVSTPP--TTASVPETTAPPAAVPTEPIDVSVLSEAAIETPVAPPVEVTTEVAVADV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAESIPAPVVATTPVPAT---LAVTDPDVTASAVPELPPVIAPSPVPSAVAETPVDLAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TPEKL-APTTPEKPAPTTPEELAPTTPEEPTPTTPEEP---APTTPKAAAPNTPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAETPEVASVAVAETTPPVVPPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTP-----KEPAPTTTKKPAPTAPKEP---APTTPKET----APTTPKKLTPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PLAAAEPVVVAPPATETPVVAPAAASPHVSVAPAVETAVVAPVSASTEPPVAAATLTTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTTTKEP---APTTTKSAPTTPKEPSPTTTKEPA-PTTPKEPAPTTPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAVPAAVPVVAPVLAPAVAPAVAPVVAETPAPPPVAETPVATTPECVAPLTPEVSVVATK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKEPA--PTAPKKPAPTTPKEPAPTTPKEPAPTTTKE-PSPTTPKEPAP----TTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PVAATPTEVVQIPVAAPVIATEPVAASAPTPAAVTPVISPVIASPPVVPANTTVPVAAPV
                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                    Mytilus.
                                            Yamaha E.,
                                                                                                                                                                                                                                                                      872
               mussel
                                                                                                                                                                                                                                                                      ₹
              Mytilus
                                          Yamazaki F.,
                                                                                                                                                                                            PROTEIN
            coruscus and
                                                                                                                                                                                            ۳
                                                                                                                                              Mytiloida;
                                                                                                                                                                                                                                                                                                                                             864
                                          Odo
                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              387
```

```
Ş
                                                                                    망
                                                                                                                                    Qγ
                                                                                                                                                                                    B
                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                       ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SET TETT TO DESCRIPTION OF THE PROPERTY OF THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
Matches 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Signal;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Mol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
DOMAIN
                                     457
                                                                                    411
                                                                                                                                                                                    351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                       402
                                                                                                                                                                                                                                   352
                                                                                                                                                                                                                                                                                        293
                                                                                                                                                                                                                                                                                                                                        297
                                                                                                                                                                                                                                                                                                                                                                                        237
                                                                                                                                                                                                                                                                                                                                                                                                                                       244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D63777; BAA09850.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MO1. EVO1. 43:348-356(1996).

FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

SUBCELLULAR LOCATION: SECRETED.

TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

DOMAIN: ALMOST EXCLUSIVELY COMPOSED OF REPEATS OF A DECAPEPTIDE. PTM: PROLINES IN THE REPEAT ARE HYDROXYLATED (SINGLE OR DI-) AND ALSO TYROSINE (THUS PRODUCING DOPA = 3,4-DIHYDROXYPHENYLALANINE).
                                                                       T-YKPKISYPPTYKPKASYVSSYKSKKTYPPTYKPKISYPPTYKPKPSYPPTYKPKITYP
PT-PTTPEEPAPTTPKAAAPNTPK-EPAPTTPKEPAPTTP-----KEPAPTT--PKETAP
                                                                                                                       APTAPKEPAPTTPKETAP-----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEE
                                                                                                                                                                         TYKPKTTYPPTYKPKPSYPPSYKPKKTYSPTYKPKITYPPTYKPKPSYPPSYKPKTTYPP
                                                                                                                                                                                                                   EPSPTTTKEPA-PTTPKEPAPTTPKKP-APT-TPK-----EPAPTTPKEPAPTTTKKP
                                                                                                                                                                                                                                                                        K-PKPSYPPTYKPKITYPPT-YKPKPSYPTPYKQKPSYPPIYKSKSSYPTSYKSKKTYPP
                                                                                                                                                                                                                                                                                                                     KEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPT----TTKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                        SYPSIYKPKASYVSSYKSKKTYPPTYKPKISYPPTYKPKPSYP----PTYKPKVTYPPTY
                                                                                                                                                                                                                                                                                                                                                                                                       PAPTTPKEPAPTT----PKEPAPTT--PKEPAPTTPKEPAPT-APKKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RKPSYTPYKPKATYPPTYKPKITYP----PTYKRKPSYT----PYKPKTTYPPTYKPKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YPPTYK-----PKITYP--PTYKQKPSYPPSYKPKTTYPPTYK----PKITYPPTYK 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SAPT-TPKEPAPT--TTKEPAPTTPKEPAPTTTKSAPTTPKEPAPTT--PKK 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AKPTPKAETTTKGPALTTPKEPT-PTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTK 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YPTKRYQPTYGSKTNYPPIYKPIAKKLSSYKAIKTTYPAYKAKTSYPPSYK-----HKIT 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VVDEAGSGLDNGDFKVTTPDTST---TQH-----NKVSTSPKITTAKPINPRPSLPPNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VYGSAYSGASAGAYK-TLPGSHPYGSKHVPVYKPMNKIPT-PYI--SKKSYPAPYKPKGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       o; IPR002964; Adhesive_plaq.
o; IPR002965; P_rich_extensn.
PR01216; ADHESIVEI.
PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   184
213
872 /
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
21
124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ã,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydroxylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   872
41
872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           192
221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΨW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 518.5; [
Pred. No. 3.5e:
80; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NONAPEPTIDE 1.
NONAPEPTIDE 2.
NO 98CC70D7C75FF3C4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ADHESIVE PLAQUE MATRIX PROTEIN MONREPETIVE LINKER.
TANDEM REPEATS OF Y-K-[PS]-K-[P-[PST]-[ST].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF Y-K-[PS]-K-[IP]-[ST]-Y-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 872;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                          507
                                                                                                                         456
                                                                                                                                                                                                                         401
                                                                                                                                                                                                                                                                        350
                                                                                                                                                                                                                                                                                                                           351
                                                                                                                                                                                                                                                                                                                                                                                                                          296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     69
                                                                                                                                                                                                                                                                                                                                                                             292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54;
```

```
RESULT 10
CPN_DROME
ID CPN_D
AC Q0291
DT 01-0C
DT 01-0C
DT 01-FE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                          дЬ
                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                     Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                          Qγ
                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                        망
                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                밁
                                                                                                                                                                     δÃ
                                                                                                                                                                                             B
                                                                                                                      QΥ
                                                                                                     멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
REPEAT
CPN_DROME
Q02910;
Q1-OCT-1993
Q1-OCT-1993
Q1-FEB-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02012; BNR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Phosphorylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                      1126
                                                                                                                                                                                              1034
                                                                                                                                                                                                                                                                                                          478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             599
                                                                                                                           712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      211
                                                                                                                                                                                                                     597
                                                                                                                                                                                                                                          984
                                                                                                                                                                                                                                                                538
                                                                                                                                                                                                                                                                                     936
                                                                                                                                                                                                                                                                                                                                877
                                                                                                                                                                                                                                                                                                                                                      428
                                                                                                                                                                                                                                                                                                                                                                             818
                                                                                                                                                                                                                                                                                                                                                                                                   373
                                                                                                                                                                                                                                                                                                                                                                                                                        765
                                                                                                                                                                                                                                                                                                                                                                                                                                             317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KE-PAPTTTKEPAPTTTKSAPTTPKE-PAPTTPKKP-----APTTPKEPA----PTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPSTPADSSAHSTPSTPVDS---SAHSTPSTPADSSAHGTPSTPVDSSAHGT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPTTPKE-PASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTKEPAPTTP 210
                                                                                                                                                                                                                                                                                                        PKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT
                                                                                                                                                                                                                                                                                                                                                     TTP-EKLAPTTPEKPAPTTPEELAPTTPEE----PTPTTPEEPA----PTTP-KAAAPNT
                                                                                                                                                                                                                                                                                                                                                                           {\tt TPVDSSAHSTPSTPVDSSAHGTPSTPVDSSAHSTPSTPADSSAHST-PSTPADSSAHGTP}
                                                                                                                                                                                                                                                                                                                                                                                                 TP-KKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKL----TP
                                                                                                                                                                                                                                                                                                                                                                                                                                            PTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VDSSAHGTPSTPADSSAHSTPSTPADSS-----AHSTPSTPADSSAHSTPSTPVDSSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KEPTP-TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      al Similarity
177; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P29768;
                                                                                                      GT
                                                                                                                                                                                                                  KPAPKELAPTTTKGPTSTTSDKPAPTTPK----ETAPTTPKEPAPTTPKKPAPTTPETPP
                                                                                                                                                                                                                                                                                    PSTPADSS----AHSTPSTPADSSAHST-PSTPADSS-----AHSTPSTPVDSSAHST
                                                                                                                                                                                                                                                                                                                                 STPVDSSAHSTPSTPADSSAHS-TPSTPVDSSAHSTPSTPADSSAHGTPSTPVDSSAHGT
                                                                                                                                                                                                                                                                                                                                                                                                                         HGTPSTPADSSAHSTPST---PADSSAHGTPSTPVDSSAHST----PSTPVDSSAHGTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADSSAHGTPSTPVDSSAHSTPSTPVDSSAHSTPSTPVDSSAHGAPSTPADSSAHGTPSTP
                                                                                                                           M
                                                                                                                                                DSSAH-STPSTPAD-SSAHGTP--STPADSSAHSTP
                                                                                                                                                                                                                                                                TTKEPTSTTSDKPAPTTPKGTAPTTPKE-PAPTTPKEPAPTTPKGTAPTTLKEPAPTTPK
                                                                                                                                                                      PTTSEVSTPTTTKEPTTIHKSPDESTP-ELSAEPTPKALENSPKEPGVPTTKTPAATKPE
                                                                                                                                                                                              TP----ADSSAHGTPSTPADSSAHSTPSTPVDSSAHSTPSTPADSS----
                                                                                                      1127
                                                                                                                           713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            768; 1DIL.
IPR002860;
                                                                                                                                                                                                                                         -PSTPADSSAHST----PSTPADSSAHSTPSTPADSSAHSTPSTPVDSSAHSTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Glycosidase; Glycoprotein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23
163
209
458
589
342
394
1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
  (Rel.
(Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
34
174
220
588
1120
342
394
1125
   27,
27,
28,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                120032
  Last
Last
                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BNR 1.
BNR 2.
BNR 3.
FIBRONECTIN TYPE-III.
44 X 12 AA TANDEM REPEATS, L'
44 X 12 AL TANDEM REPEATS, L'
45 YOUR REPEATS, L'
46 YOUR REPEATS, L'
47 YOUR REPEATS, L'
48 YOUR REPEATS, L'
49 YOUR REPEATS, L'
49 YOUR REPEATS, L'
40 YOUR REPEATS, L'
40 YOUR REPEATS, L'
40 YOUR REPEATS, L'
41 YOUR REPEATS, L'
42 WW; 07049221897C6A40 CRC64;
   sequence u
annotation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CYS-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 e 530.5; DB 1,
No. 1.5e-14;
             update)
                                                  865
   update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GPI-anchor;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LTR DOMAIN
                                                                                                                                                      STPAGSSAN 1125
                                                                                                                                                                                                 AHGTPSTPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -PSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                   -PT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                    596
                                                                                                                                                                                                                                                                                                                                                                                                       427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                651
                                                                                                                                                                                                   1084
                                                                                                                                                                                                                                              1033
                                                                                                                                                                                                                                                                                                                                                                                876
                                                                                                                                                                                                                                                                                                                                                                                                                                                  372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         764
                                                                                                                                                                                                                                                                                                                                                           477
                                                                                                                                                                                                                                                                                                                537
                                                                                                                                                                                                                                                                                                                                      935
                                                                                                                                                                             711
                                                                                                                                                                                                                       652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34;
```

```
QY
                                                                                 Вþ
                                                                                                      Qy
Q
                           망
                                                                                                                                                 Query Match
Best Local :
                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila
Eukaryota;
                                                                                                                                                                                                        FlyBase; FB90001
Calcium-binding.
Calcium-binding.
CONFLICT 36
CONFLICT 64
CONFLICT 76
CONFLICT 100
CONFLICT 154
CONFLICT 154
CONFLICT 534
CONFLICT 530
CONFLICT 59
CONFLICT 703
CONFLICT 721
SEQUENCE 865 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Calcium and contains a leucine zipper.";

Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).

Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).

Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).

REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT

OF CA+2 PER MOL OF PROTEIN.

OF CA+2 PER MOL OF PROTEIN.

OF CA+2 PER MOLOMER (PROBABLE).

-1- SUBCULLULAR LOCATION: CYTOPLASMIC; HYPODENSE COMPART

-1- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRALN=CANTON-S;
MEDILINE=93165729; PubMed=8094559;
Martin J.H., Benzer S., Rudnicka M., Miller C.A.;
"Calphotin: a Drosophila photoreceptor cell calcium-binding proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Arthropoda; Tracheata; He
Pterygota; Neoptera; Endopterygota; Diptera;
Ephydroidea; Drosophilidae; Drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CALPHOTIN.
CPN OR CAP.
                                                                                                                                                                                                      CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ephydroidea; Dro
NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; L02111; AAA28
EMBL; L05080; AAA28
PIR; A47282; A47282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ballinger D.G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=93165730;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-CANTON-S;
                                                                                                         123 APTSKVLAKP-TPK---AETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTT
                                                                                                                                      Local Similarity
nes 233; Conser
                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         prosophila photoreceptor cell-specific
                                                                                 œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPOUND EYES AND OCELLI.
DEVELOPMENTAL STAGE: EXPRESSED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEVELOPMENT.
                            VTIPAPAPIAAASVTPVASVAPPVVAAPTP-----PAASPVSTPVAVAQIPVAVSAPVAP
                                                                                 SPVSAPVAAPVTPSAVAAPVQVVSPAAVAP-----APAAPIAVTPVAPPPTLASVQPAT
  KEPAPTTPKKPAP-
                                                                                                                                                                                                                                                                                                                                                                                                FBgn0010218; Cpn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       melanogaster (Fruit fly).
                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                         AAA28405.1;
AAA28420.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Xue
                                                                                                                                                                                                   36
43
64
76
100
127
154
154
160
534
699
703
701
84781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PubMed=8434015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N., Harshman K.D.;
                                                                                                                                                    10.6%;
26.2%;
    -TTPKEPAPTTPKEPTP-----TTPKEPAPTTKEPAPTT
                                                                                                                                                                                                           WW;
                                                                                                                                       73;
                                                                                                                                                                                                                   A -> AVAPAVVA (IN REF. 2)
I -> Y (IN REF. 2)
I -> V (IN REF. 2)
T -> A (IN REF. 2)
P -> PP (IN REF. 2)
VQ -> AP (IN REF. 2)
I -> V (IN REF. 2)
S -> T (IN REF. 2)
A -> E (IN REF. 2)
I -> T (IN REF. 2)
O -> L (IN REF. 2)
V -> L (IN REF. 2)
                                                                                                                                      Score 530; DB
Pred. No. 1.2e
73; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Tracheata; Hexapoda; Ins
ota; Diptera; Brachycera;
                                                                                                                                                                                                           2110417E0B0E7CFE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EARLY IN
                                                                                                                                                      530;
No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HYPODENSE COMPARTMENT.
                                                                                                                                                      DB 1;
.2e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein,
                                                                                                                                          433;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PHOTORECEPTOR
                                                                                                                                                                   Length
                                                                                                                                                                                                            CRC64;
                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              calphotin,
                                                                                                                                                                                                                                                                                                                                                                         2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Insecta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "SPONGE"
T BINDS 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLS
                                                            -TTKSAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Muscomorpha;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELL
                                                                                                                                            152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 binds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .3 TO
                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOL
                                                                                                                  178
        279
                                   116
                                                             233
                                                                                        61
                                                                                                                                             41;
```

```
Š
                                                                                   В
                                                                                                                                                               밁
                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                  Вb
                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DR CREATE STATES OF STATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
DOMAIN
          905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Repeat.
                                               506
                                                                                                                        458
                                                                                                                                                                                                                                                                                                                      677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                               790
                                                                                                                                                                                                      398
                                                                                                                                                                                                                                         730
                                                                                                                                                                                                                                                                                343
                                                                                                                                                                                                                                                                                                                                                         291
                                                                                                                                                                                                                                                                                                                                                                                                620
                                                                                                                                                                                                                                                                                                                                                                                                                                     246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pram; PF00038;
PROSITE; PS002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Intermediate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 DEAGSGLDNGDFKVT-----TPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKET
TLPTPKTEAKESKKDEAPKEAPKPKVEEKKETPTEKPKDSTAEAKKEEAGEKKKAVASEE
                                      APTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSD--KPAPTTPKGTAPTTP
                                                                                                            TPTTPEE-PAPTTPKAAAPNTPKEPAPTT----PKEPAPTTP------
                                                                                                                                                                                          TKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEP
                                                                                                                                                     AKSPVKEDIKPPAEAKSPEKA-KSPVKEGAKPPEKAKPLDVKSPEAQTPVQEEATVPTDI
                                                                                                                                                                                                                                 SPAAVKSPGEAKSPGEAKSPAEAKSPAEAKSPIEVKSPEKAKTPVKEGAKSPAEAKSPEK
                                                                                                                                                                                                                                                                                                                                             P----APTTPKEPAPTTPKEPAPTTKEPSPTTPKEPA----PTTTKSAPTTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VIVEGQTEEIRVTEGVTEEEDKEAQGQEGEEAEEGEEKEEEELAAATSPPAEEAASPEKE
                                                                                                                                                                                                                                                                     TKSAPTTPKE-PSPTTTKEPA----PTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                          PGEAKSPAEPKSPAEA--KSPA----EVKSPAEAKSPAEVKSPGEAKS-PAAVKSPAEAK
                                                                                                                                                                                                                                                                                                                                                                                        AKSPAEAKSPAEAKSPATYKSPGEAKSPSEAKSPA---EAKSPAEAKSPAEAKSPAEVKS
                                                                                                                                                                                                                                                                                                                                                                                                                             ---PTTPKEPA----PTTPKEP----TPTTPKEPAPTTKEPAPTTPKEPA----PTAPKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TKSRVKEEAKSPGEAKSPGEAKSPA----EAKSPGEAKS-PGEAKSPGEAKSPAEPKSPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EPKSPAEAKSPA--EPKSPA--TVKSPGEAKSPSEAKS-PAEAKSPAEAKSPAEAKSPAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EP-APTTTKEPAPTTPKEPAPTTTKEP----APTTTKSAPTTPKEPA----PTTPKKPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SLTVNKETTVETKETTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPK 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EECRIGFGPSPFSLTEGLPKIPSIST - - HIKVKSEEMIKVVE-
                                                                           - PEQVKSPAKEKAKSPE - - KEEAKTSEKVAPKKEEVKSPVKEEVKAKEPPKKVEEEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1087
1087
1087
1087
1087
1087
1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         filament;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; filament;
226; IF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  408
1087
817
886
1087
129
141
239
241
281
281
199
281
199
281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Coiled coil; Neurone;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COIL 2B.

K -> QA (IN REF. 2 AND 3).
A -> AR (IN REF. 2 AND 3).
S -> T (IN REF. 2 AND 3).
L -> G (IN REF. 2 AND 3).
P -> PREAKSP (IN REF. 3).
MISSING (IN REF. 3).
G -> A (IN REF. 3).
G -> A (IN REF. 3).
V -> M (IN REF. 2 AND 3).
T -> N (IN REF. 2 AND 3).
T -> N (IN REF. 2 AND 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKER 1.
COIL 1B.
TUKER 12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 533; DB 1;
Pred. No. 1.1e-14;
1; Mismatches 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINKER 12
COIL 2A.
LINKER 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLU-RICH (ACIDIC).
50 X 6 AA TANDEM REPEATS
GLU/LYS-RICH.
COIL 1A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1087;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------KSEKET
                                                                                                                KEPAPTTPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    K-S-P-A-E-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
  964
                                                                           904
                                                                                                                                                       848
                                                                                                                                                                                                                                   789
                                                                                                                                                                                                                                                                       397
                                                                                                                                                                                                                                                                                                                                                                                        676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38;
```

Ş

Ş

```
RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                RX: MEDLINE-91376547; Pubmed-1896773;
RA Prioli R.P., Mejia J.S., Aji T., Aikawa M., Pereira M.E.A.;
"Trypanosoma cruzi: localization of neuraminidase on the surface of trypanosoma cruzi: localization of neuraminidase on the surface of trypanosoma cruzi: localization of neuraminidase implicated in confiction. Developmentally regulated Neuraminidase implicated in confidency of cells.

C -1- FUNCTION: DEVELOPMENTALLY REGULATED NEURAMINIDASE IMPLICATED IN CONTROL ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC COLTANAGES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEGRAMINYL RESIDUES TO GALACTOCE, N-ACETYLHEXOSAMINE, OR N-OR O-ACYLATED CONTROL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS, GLYCOLIPIDS OR COLOMINIC ACID.

C SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPT-ANCUAD CONTROL ACID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
STRAIN-SILVIO X-10/4;
MEDLINE-91277609; PubMe
Pereira M.E.A., Mejia J
EMBL;
PIR; ;
                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                              <del>-</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "The Trypanosoma cruzi neuraminidase contains bacterial neuraminidases, YMTD repeats of the receptor, and type III modules of fibronectin.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1991 (Rel. 20, Created)
01-NOV-1991 (Rel. 20, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
SIALIDASE (EC 3.2.1.18) (NEURAMINIDASE) (NA) (MAJOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBCELLULAR LOCATION MEDLINE-91376547; Pul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Euglenozoa;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Prioli R.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=5693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Trypanosoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TCNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TCNA_TRYCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     624
                                                                                                                                                                                                                    DEVELOPMENTAL STAGE: MAXIMAL ACTIVITY IN TRYPOMASTIGOTES, IN EPIMASTIGOTES AND NO DETECTABLE IN AMASTIGOTES.

MISCELLANEOUS: THE VARIABLE LENGTHS OF THE LONG TANDEM RE
DOMAIN COULD ACCOUNT IN PART FOR THE POLYMORPHISM OF THE
                                                                                                                                                                          SIMILARITY:
SIMILARITY:
                                                                                                                                                                                                                                                                                            (POSSIBLE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TKGEK 1087
               м61732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TTTEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ---PK-MEAKVKEDDKSLSKEP--SKPKTEKAEKSSSTDQKESQPPE-----KTTEDKA 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ETPAKLGVKEEA--KPKEKTETTKTEAEDTKAKEPS----KPTETEKP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cruzi.
 : AAA30255.1;
JH0557.
                                                                                                                                                                          BELONGS TO FAMILY 33 OF CONTAINS 3 BNR REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PubMed=1711561;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kinetoplastida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ortega-Barria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Trypanosomatidae;
                                                                                                                                                                                          GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      E., Matzilevich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KDTKEEKTTESRKPEEK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequences similar to
low density lipoprotein
.";
                                                                                                                                                                                          HYDROLASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SURFACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Trypanosoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ANTIGEN).
                                                                                                                                                                                                                                                                        MINIMUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1034
```

뫄

```
Query Match
Best Local S
Matches 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           entities requires a license agreement (See http://www.isb-sib.ch/announce,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             use by non-profit institutions as long modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the European Bioinformatics Institute.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         136
                                                                                                                                                                                                                                                                                                                                                                                                      346
                                                                                                                                                                                                                                                                                                                                                                                                                                        194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        185
                                                                                                482
                                                                                                                                                                   428
                                                                                                                                                                                                   507
                                                                                                                                                                                                                                      370
                                                                                                                                                                                                                                                                    450
                                                                                                                                                                                                                                                                                                      310
                                                                                                                                                                                                                                                                                                                                     400
                                                                                                                                                                                                                                                                                                                                                                       250
685
                                                                                                                                  567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTM: EXTENSINS C
SER-PRO(4). THE
GLYCOSYLATED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SWISS-PROT entry is copyright. It is produced through en the Swiss Institute of Bioinformatics and the E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AETTTK-GPALTTP--KEPTPTT-----PKEPASTTPKE----PTPTTIKSAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAPTTPKEPAPTTPKEPAPTTTKEPSPTTPK---EPAPTTTKSAPTTTKEPAPTTTKS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTTKSAPTTPKEP----APTTTKEPAPT---TPKEPAPTTTKEPAPTTTKSAPTTPKE
                                                                                                                                                                                                                                                                                                      PPTPTPTFSPPPPAYSPPPTYSPPPPTYLPLPSSP1YSPPPPVYSPPPPPSYSPPPPTYL
                                                                                                                                                                                                                                                                                                                                                                       RHLPPSPRRQPQPPTYSPPPPAYAQSPQPSPTYSPPPPTYSPPPPSPTYSPPPPAYSPSP
                                                                                                                                                                                                                                                                                                                                                                                                     ---APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPT---TPKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-----PSPSHGHAPPSGGHTPPRGQHPPSHRRPSPPSRHGHPPPPTYAQPPPTPIYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAPTTPKKPAPT - - -TPKEPAPTTPKEPTPTTPKEPAPTTK - -EPAPTTPKEPAPTAPKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SHGHLPPSVGGPPPHRGHLPPSRGFNPPPSPVISPSHPPPSYGAPPPSHGPGHLPSHGQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AEATTQYGGYLPPPVTSQPPPSSIGLSPPSAPTTTPPSRGHVPSP---RHAPPRHAYPPP
                                                                                                                                  APTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKE
                                                                                                                                                                                                                                    PPPPPSSPPPPSFSPPPPTYEQSPPPPPAYSPPLPAPPTYSPPPPT--YSPPPPTYAQPP
                                                                                                                                                                                                                                                                 APTTPEEPTPTTPEEPAPTTPKAAAP----NTPKEPAPTTPKEPAPTTPKEPAPTTPKETA 506
                                                                                                                                                                                                                                                                                                                                       KPAPT - - - APKEPAPTTPKETAPTTPKKL - TPTT - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                        PSPQV-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           170;
PTPKALENSPKEPGVPTTKTPAATKP
                                 PPPRQIHSPPPPHWQPRTPTPTYGQPPSPPTFSAPPPRQIHSPPPPHRQPRPPTPTYGQP
                                                               TAPTTPKEPAPT - - TPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                PPPSPIYSPPPPQVQPLPPTFSPPPPRRIHLPPPPHRQP--
                                                                                                                                                                   P----
                                                                                                                                                                                                   PTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X13885; CAA32090.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cell wall;
                                                                                                                                                                                                                                                                                                                                                                                                                                        -QPPPT--YSPPPPTHVQPTPSPPSRGHQPQPPTHRHAPPTHRHAPPTHQPSPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                   LPPTYSPPPAYSPPPPTTYSPPPPTYSPPPAYAQPPPPPPTYSPPPPAYSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S06733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE E PROLINE RESIDUE IS HYDROXYLATED AND THEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            600
65406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          620
73
151
242
235
242
620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11.0%;
27.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EXTENSIN.
H-A-P-P.
H-A-P-P.
2 X 7 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 551;
Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONTAINS THE SER-PRO(4)
3 X APPROXIMATE TANDEM I
641DD2278AB28524 CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TANDEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1
1.5e-15;
304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                There are no rest
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Structural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Usage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEATS
                                                                                                                                                                                                                                                                                                                                       -PEKLAPTTPEKPAPTTPEEL
                                                                                                    -RPPTPTYGQPPSPPTFSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Уď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T-H-R-H-A-P-P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      commercia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             184
                                                                                                                                                                                                                                                                                                                                           449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             193
                                                                                                                                                                                                                                          427
                                                                                                                                                                                                                                                                                                                                                                            309
                                                                                                                                                                                                                                                                                                                                                                                                                                           249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24;
```

δÃ 밁 QΥ В Ωy

QY B δÃ 밁

Qy В δõ 밁 Š 망 δÃ 밁 δõ ₽ δõ В

EMBL;

PIR; PIR; MGD;

A43778; A43778. MGI:97309; Nfh.

QFMSH

IPR001664;

ΙF

```
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-NOV-1990 (Rel. 16, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT
(NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NFH_MOUSE
P19246; Q61959;
                                                                                                                                                                                                                                                                                                                                                                              SUBMITTED (MAR-1994) to the EMBL/GenBank/DDBJ databases.

-I- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS:
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEUROWAL CALI
NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT
SUBSERVED BY THE TWO SMALLER NF PROTEINS.

-I- PTM: THERE ARE A NUMBER OF THE SERINES IN THIS MOTIF. IT
PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THE FORMATION
THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Shneidman P.S., Carden M.J., Lees J.F., Lazzarini F "The structure of the largest murine neurofilament revealed by cDNA and genomic sequences."; Brain Res. 464:217-231(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Sequence and structure neurofilament subunit."; Gene 68:307-314(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Julien J.-P.,
Mushynski W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. MEDLINE=89121513; Pubmed=3220257;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NEFH OR NFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10090;
                                                                                            EMBL;
                                                                                                                                                            entities
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-SWISS WEBSTER;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=89089138; PubMed=3145094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      599
                                                                                                                      EMBL;
                                                                                                                                                  or send an
                                                                                                                                                                            use by non-profit institutions as long modified and this statement is not removed
                                                                                                                                                                                                                                  This
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Carden M.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                      PIM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCT
OF THE LARGER NEUROPILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
LEVELS OF PHOSPHORYLATION BEING ALTERED EXCELOPMENTALLY AND
COINCIDENT WITH A CHANGE IN THE NEUROPILAMENT FUNCTION.
SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITION:
                                                                                                                                                                                                    European
                                                                                                                                                                                                                                                                                                                                                     INTERFILAMENT CROSS BRIDGES THAT ARE OF AXONAL CALIBER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PSP-
                                                 , M24496; AAA39813.1; M23349; AAA39813.1; J
, M24494; AAA39813.1; J
, M24494; AAA39813.1; J
, M24495; AAA39813.1; J
, M35131; AAA39809.1; A
, Z31012; CAA83229.1;
                                                                                                                                                                                                   SWISS-PROT entry is copyright. It is produced through een the Swiss Institute of Bioinformatics and the Eleropean Bioinformatics Institute. There are no rest
                                                                                                                                                                                                                                                          TO 716
                                                                                                                                                 requires a license agreement (See an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                          AND IS SHORTER DUE TO FRAMESHIFTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Cote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chordata;
Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTTYSPPSPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Beaudet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of
                                                    JOINED.
JOINED.
ALT_FRAME.
                                                                                                          JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Craniata; Vertebrata;
Sciurognathi; Muridae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sidky
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                            There are no rest
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B
                                                                                                                                                                                                                                                                                                                                                                       IMPORTANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      coding
                                                                                                                                                                 http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Muridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Flavell
                                                                                                                                                                               Usage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        for
                                                                                                                                                                                                                                                                                                                                                                                                                                       E IF PROTEINS: L, M, F NEURONAL CALIBER. S THAT IS NOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R.A.;
                                                                                                                                                                                                                                                                                                                                                                        H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Euteleostom1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Murinae;
                                                                                                                                                                                                                                                                                                                                                                         THE MAINTENANCE
                                                                                                                                                                                                          restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTEIN)
                                                                                                                                                                                                                                                                             POSITIONS
                                                                                                                                                                                                                                                                                                                                THE FUNCTIONING NF-H), THE
                                                                                                                                                                                                                                                                                                                                                                                      FORMATION
                                                                                                                                                                                  and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Grosveld
                                                                                                                                                                                                                          EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (NF-H)
                                                                                                                                                                                                                                      a collaboration
                                                                                                                                                                                  for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus
                                                                                                                                                                                                                          outstation
                                                                                                                                                                                                                                                                                                                                                                                                    SI
                                                                                                                                                                                                                                                                                                                                                                                        OF
                                                                                                                                                                                                                                                                              534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ĭ
                                                                                                                                                                                                           its
```

```
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                 Pfam; Prva.

Pfam; PF00094; Vwu,

Pfam; PF000917; PRICHEXTEIN.

PRINTS; PR01217; MAM; 2.

SMART; SM00117; WAC; 1.

SMART; SM00011; VWC; def; 3.

SMART; SM00011; VWD; 3;
                                         CONFLICT
NON_TER
SEQUENCE
                                                                         CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                    DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                             PROSITE; PS01186; EGF_2;
PROSITE; PS00740; MAM_1;
PROSITE; PS50060; MAM_2;
                                                                                                                                                                                DOMAIN
                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                     Ptam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MIM; 602372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS. THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS. DOMAIN: THE VWED DOMAINS 2 AND 3 MAY MEDIATE COVALENT OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2). SIMILARITY: CONTAINS 3.5 MAM DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS
DOMAIN: THE MAN DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZONA PELLUCIDA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF053356;
U83191; AA
                                                                                                                                                                                                                                                                                                                                                                                                                                            PF00629; MAM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR003328; IPR001007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR00291
                                                              1023
1099
1618
1737
1737
1832
1878
2136
2374
                                                                                                                                                                                        953
1066
1455
1862
2293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6; AAC78790.1; -. AAC51208.1; -.
                                                                                                                                                                                                                                                                                                                    ransmembrane;
                                          3
                                                                                                                                                                                                                                                                                                                                                                                                    PRICHEXTENSN
                                                                                   74
403
1023
1099
1618
1737
1832
1832
2136
  11.1%; 32.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ۷wd
                                         293013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EGF-like.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            _rich_extensn.
                                         ¥.
                                                                                                                                                                                       (MUCIN-LIKE DOMAIN).
VWED 1 (PARTIAL).
VWED 2.
VWED 3.
VWED 4.
VWED 5.
 Score
Pred.
                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                         N-LINKED
                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                   Cell adhesion;
                                                                                                                                                                                                                                                   MAM 4.
66 X HEPTAPEPTIDE REPEATS
                                                                                                                                                                   N-LINKED
                                       80E60CC0B12277B1
No.
                                                                                                                                                                                                                                                                                   (PARTIAL).
                                                                      (GLCNAC
.5; DB
3e-15;
                                                             RAGPGP
         1;
        Length 2700
                                       CRC64;
                                                                                        (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                      (POTENTIAL)
                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                   (APPROXIMATE)
```

```
RESULT 7
EXTN_TOBAC
ID EXTL_T
AC P13983
DT 01-JAN
DT 01-JAN
DT 01-JAN
DT 01-GEN
CO SPERMA
OC SPERMA
OC ASTECT
RN [1]
RP SEQUEN
RC STRAIN
RX MEDLIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             뮹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                        "Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation."; Genes Dev. 3:1639-1646(1989).
-!- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EXTN_TOBAC STANDARD; PRT; 620 AA.
P13983;
01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-MAR-1992 (Rel. 21, Last annotation update)
EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN=CV. XANTHI; TISSUE=Leaf;
MEDLINE=90128263; PubMed=261290
                                                                                                                                                                                                                                                                                                                                      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheo Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
                                                                                                                                                                                        Keller B., Lamb C.
                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=4097;
                                                                                                                                                                                                                                                                                                                                                                                                                                Nicotiana tabacum
                                                                                                                                                                                                                                                                                                                                                                                                                                                              HRGPNT3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     912
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58
                         THE MECHANICAL PENETRATION OF THE CORTE MAIN ROOT.
  SUBCELLULAR LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TLKEPAPTTPKKPAPKELAPTTTKEPTSTT-----SDKPAPTTPKGTAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PASCKSPRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTTPKEPAPT 638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PPHPSPTA-TGLAALVMSPHAPSTPMTSV---ILGTTTTSRSSTGMSCP-PNARYESCAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKET--A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --EKPTISPEKLTIPTEKPTISTEKPTIPTEKLTIPTEKPTIPTEKPTIPT--EKLTALR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KPTISPEKPTISTEK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EKPTISTEKPTIPTEKPTISPEKLTIPT-EKLTIPTEKPTIPIEETTISTEKL--TIPTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEPAPTTPKETAPT-----TPKKLTPTTPEKLA-----PTTPEKPAPTTPEELAPTTPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EKPTISTEE--PTTPTEETTISTEKPSIPMEKPTLPTEETTTSVEETTISTEKLTIPM--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EPSPTTTKEPAPTTPKEPAPTTPKKPA----PTTPKEPAPTTPKEPAPTTTKKPAPTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --EKPSIPTEK---PTISMEETIISTEKPTICPEKPTIPTEK---PTIPTEKSTISPEK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TEKPTIPSE-KPTIPSEKPTISTEKPTVPTEE---PTTPTEETTTYMEEPVIPTEKPSIPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -PTTPTE-KPTIPTEKPTISTEKPTIPTEK-PTISPEKLTIPTEKLTIPT---EKPTIPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTTKEPA-PTTTKSAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SVTTEKPT-----VPKEKPTIPTEKPTISTEK----PTIPSEKPNMPSEKPTIPSEKPTIL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPT-----PTTI 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PVKVLPELPPVSPVSS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           976
                                                                                                                                                                                                                                                                                                                                                                                                                                (Common
                                                                                                                                                                                                                 PubMed=2612909;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -TGPSETTGLTENPTISTK-------KPTVSIEKP
EXTRACELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                tobacco)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTIPTE-KPTIPTE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                  CORTEX AND
                                         FUNCTION, POSSIE EX AND EPIDERMIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ETTISTEKLTIPT
                                                                                                                                                                                                                                                                                                                                                                                              Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        107;
                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32;
```

```
Qy
                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                            망
                                                                                                                                                                                                                                                                                                                             20
                                                                                                                                                                                                                                                                                                                                                                    B
                                                                                                                                                                                                                                                                                                                                                                                                 ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                           ₽
                         δÃ
                                                                     Вp
                                                                                                    Ωy
                                                                                                                                             ₽
                                                                                                                                                                                                                      밁
맑
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein.";
Biotechnol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hydroxyproline-containing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Waite J.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PARTIAL SEQUENCE,
MEDLINE=83135732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the European Bioinformatics Institute.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       + + + +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        J. Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mussel,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Evidence for a repeating
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X54422; CAA38294.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                118
                                                                                                                                                                                                                        178
                                                                                                                                                                                                                                                                                                                                 143
 354
                                  340
                                                                         294
                                                                                                            290
                                                                                                                                                 234
                                                                                                                                                                                  241
                                                                                                                                                                                                                                                           194
                                                                                                                                                                                                                                                                                                                                                                      60
                                                                                                                                                                                                                                                                                                                                                                                                        88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sel, Mytilus edulis L.";

Biol. Chem. 258:2911-2915(1983).

FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

SUBCELLULAR LOCATION: SECRETED.

TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

DOMALN: ALMOST EXCLUSIVELY COMPOSED OF REPEATS OF A DECAPEPTIDE. DOMALN: ALMOST EXCLUSIVELY COMPOSED OF REPEATS OF A DECAPEPTIDE. PTM: THE DECAPEPTIDE A K-P-S'-P-T-Y-K IS POST-TRANSLATIONALLY MODIFIED AS FOLLOWING: THE SIXTH AND SEVENTH RESIDUES ARE HYDROXYLATED AND THE PENULTIMATE IS A 3,4-DIHYDROXYPHENYLALANINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (DOPA) DERIVED FROM TYROSINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SWISS-PROT entry is copyright. It is produced through a
sen the Swiss Institute of Bioinformatics and the EMBL
European Bioinformatics Institute. There are no restric
                                                                                                                                                                                                                                                                                                                                                                                                     KETTTTNKQTSTDG--KEKTT---SAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKG
                                                                                                                                                                                                                                                                                                                                                                                                                                               TKHEPVYKPVKTSYSAPYKPPTYQPLKKKVDYRPTKSYPPTYG-SKTNYLPLAKKLSSYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKPSYPPTYKA
                                                                                                                                                                    PKPA--PTTPKEPA--PTTPKEPT--PTTPKEP---APTTKEPAPTTPKEPA--PTAPK
                                                                                                                                                                                                                                                                                                                                                                      PIKTTYNAKTNYPPVYKPKMTYPPTYKPKPSYPPTYKSKPTYKPKITYPPTYKAKPSY--
                                                                                          KPA--PTTPKEPA--PTTPKEPA--PTTTKEPA--PTTPKEPA-PTTTKSAPTTTKEPA-
                                                                                                                                                                                                                                                         PKEPAPTTTKEPA--PTTPKEPA-----PTTTKEPA-PTTTKSAPTTPK--EPAPTT
                                                                                                                                                                                                                                                                                              PSSYKPKKTYPPTYKPKLTYPPTYKPKPSYPPTYKPKPSYPPSYKTKKTYPSSYKAKPSY
                                                                                                                                                                                                                                                                                                                                 PALTTPKEPTPTTPKE----PASTTPKEPTPTTIKSAPTTP-----KEPAPTTTKSAPTT
   PSTYKAKPSYP---
                                  PTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPA--PTTPKEPA--PTTPKEPAP
                                                                         KPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PR01216; ADHESIVEI.
PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydroxylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002964; Adhesive_plaq.
IPR002965; P_rich_extensn
                                                                                                                                                                                                                        PTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAKPSYPPTYKAKPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Prog.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AND POST-TRANSLATIONAL PubMed=6298211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6:171-177(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100412
 - PTYKAKPSYPPTYKAKPTYKAKPSYPPTYKAKPSYPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11.1%; 28.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3,4-dihydroxyphenylalanine-
decapeptide in the adhesive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       106;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                , WM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TANDEM REPEATS OF Y-K-[PA]-K-[LP]-[ST]-Y-P-P-[ST]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6EA85312748CAACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          556.5;
No. 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..2e-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MODIFICATIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             and
protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
         409
                                          395
                                                                                                                                                      293
                                                                                                                                                                                                                              233
                                                                                                                                                                                                                                                                 240
                                                                                                                                                                                                                                                                                                    177
                                                                                                                                                                                                                                                                                                                                         193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87
                                                                               353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56;
```

δÃ

```
RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QУ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΨ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                          δã
                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                               Q9Y493; 000218;
20-AUG-2001 (Re)
20-AUG-2001 (Re)
20-AUG-2001 (Re)
20-AUG-2001 (Re)
Glockner G., Scherer S., Schattevoy Tsui L.C., Rosenthal A.;
                            SEQUENCE OF 1-2379 FROM N.A. MEDLINE=99018118; PubMed=9799793;
                                                                                        Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                        NCBI_TaxID=9606;
                                                                                                                   Homo sapiens (Human)
                                                                                                                                                    ZONADHESIN
                                                                                                                                                                                                                           ZAN_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   510
                                                                                                                                                                                                                                                                                                      827
                                                                                                                                                                                                                                                                                                                                  833
                                                                                                                                                                                                                                                                                                                                                                                             776
                                                                                                                                                                                                                                                                                                                                                                                                                           727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       566
                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KAKPSYP----PTYKAKPSYPPTYKAKPTYPSTYKAKPSYPPTYKPKISYPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KPAPTTPKETAPTTPKEPA--PTTPKKPA-PTT----PETPPPTTSEVSTPTTTKE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAPTTPKEPAPTTPKGT---APTTLKEPA--PTTPKKPA--PKELAPTTTKG-PTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTT----PKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKE----PAPTTPKETAPTT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPTYKAKPTYKAKP-----
                                                                                                                                                                                                                                                                                                                                                                                           TLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKK----P
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTTIHKSPDESTPELSAEPTPKALENSPK----EPGVPTT--KTPA----ATKPEMTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PATYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYKAKPTYPSTYKAKPSYPPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKLTYKPTYK-PKPSYPPSYKPKTTYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KPT-----YPSTYKA----KPSYPASYKAKPSYPPTYKSKSSYPSSYKPKKTYPPTYK
                                                                                                                                                                                                                                                                                                      STYKLKPSYPPTYKSKTSYPPTYNKKISYPSSYKAKTSYPPAYKPTNR
                                                                                                                                                                                                                                                                                                                                 TSTKKPKTMPRVRKPKTTPTP-
                                                                                                                                                                                                                                                                                                                                                                TYKSKSIYPSSYKPKKTYPPT - - - YKPKLTYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                           YKAKPTYKAKPTYPSTYKAKP--TYKAKPTYPPTYKAK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PPT-YKAKPSYPPTYKAKPTYKAKPTNPSTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPT
                                                                                                                                                 01 (Rel. 40, Created)
01 (Rel. 40, Last seque)
01 (Rel. 40, Last anno)
N (FRAGMENT).
                                                                                                        Metazoa;
                                                                                                                                                                                                                             STANDARD;
                                                                                           Primates;
                                                                                                        Chordata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -TYPST-YKAKPSYP----PSYKAKPSYPPTYKAKPTYKA
                                                                                                                                                                   : sequence
: annotatio
                                                                                           Craniata; Vo
Catarrhini;
                                                                                                                                                                                                                               PRT;
                                                                                                                                                                      ion
                                                                                                                                                                   on update)
                                                                                                                                                                                                                               2700
                  R.,
                                                                                                                                                                                                                                                                                                                                      RKMTSTMPELNPTSR
                                                                                                          Vertebrata; Euteleostomi;
                 Boright
                                                                                                                                                                                                                                AA
                                                                                           Hominidae;
                                                                                                                                                                                                                                                                                                                                                                                                                             ----PSYPPTYKPKPSYPP
                  Α.,
                                                                                                                                                                                                                                                                                                                                                                     -PKPSYPPSYKPKITYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -PKISYPPTYKAKPSY
                    Weber
                                                                                              HOMO
                                                                                                                                                                                                                                                                                                           874
                  ..
J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      506
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       715
                                                                                                                                                                                                                                                                                                                                                                       826
                                                                                                                                                                                                                                                                                                                                                                                                                                 777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 666
```

```
Gao Z., Harumi T., Garbers D.L.;
"Chromosome localization of the mouse zona
zonadhesin gene (ZAN).";
Genomics 41:119-122(1997).
-i- FUNCTION: BINDS IN A SPECIES-SPECIFIC
OF THE EGG. MAY BE INVOLVED IN GAMETE
                                                                                                                               analysis of 650 kb of genomic reveals 17 genes.";
                                                                                  MEDLINE=97271566; PubMed=9126492;
                                                                                            TISSUE-Testis
                                                                                                     SEQUENCE OF
                                                                                                             Genome Res.
                                                                                                                                         "Large-scale sequencing of two regions analysis of 650 kb of genomic sequence
SIGNALING.
- SUBUNIT: PRO
- SUBCELLULAR
                                                                                                     2338-2700 FROM N.A.
                                                                                                                       8:1060-1073(1998).
  PROBABLY FORMS COVALENT OLIGOMERS 
LAR LOCATION: TYPE I MEMBRANE PROT
                                                                  zonadhesin
                                                                                                                                          in huma
around
                                                                                                                                                    human
                             MANNER TO TH
  PROTEIN,
                                                                                                                                           the EPO and (
                                                                  gene
                                                                                                                                           EPO and CUTL1
   EXCLUSIVELY
                                       THE
                              HE ZONA
AND/OR
                                                                  and
                                                                   the
                                        PELLUCIDA
                                                                    human
    9
2
                                                                                                                                             Loci
```

```
Ş
       밁
                                                 밁
                                                                  Š
                                                                                      B
                                                                                                           õ
                                                                                                                                    ₽
                                                                                                                                                    Š
                                                                                                                                                                             В
                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 27.1
Matches 290; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M61170; AAB53150.1; --
EMBL; X52229; CAA36477.1; A)
EMBL; X52228; CAA36477.1; A)
EMBL; X52228; CAA369612.1; A)
EMBL; X52228; CAA78972.1; --
EMBL; Z17324; CAA78973.1; --
EMBL; Z17325; CAA78973.1; --
EMBL; S1823; AAA35757.1; --
EMBL; S81781; AAD14376.1; A
EMBL; S81786; AAD14376.1; A
EMBL; S81736; AAD14376.1; A
EMBL; S81736; AAD14369.1; A
EMBL; S81736; AAD14369.1; A
EMBL; S10218; A35175.
PIR; A35175; B35175.
PIR; B35175; B35175.
PIR; B35175; B35175.
                                                                                                                                                                                                                                                                                                        VARSPLIC
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
DOMAIN
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MIN; 158340;
MIN; 113720;
InterPro; IPR000082; SEA.
InterPro; IPR000082; SEA.
SMART; SM00200; SEA; 1.
PROSITE; PS50024; SEA; 1.
PROSITE; PS50024; SEA; 1.
Repeat; Alternative splicing.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                               VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                              VARSPLIC
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Repeat;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                    VARSPLIC
239 TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                  182
                                                                     207
                                                                                           126
                                                                                                              167
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                 134 P-KAETTIKGP--ALTIPKEP-----KE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlycoSuiteDB; P15941; -.
                                                                                                                                     66
                                                                                                                                                                                                  75
                                                                                                                                                                      TVLTV-----VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                         TSLTVNKETTVETKETTTINKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT
                    --APTTPKEPTPTTP------KEPAPTTKEP-----APTTPKEPAPTAPK----
                                                             PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP-----KKPAPTTPKEP----
                                                                                   PAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSA
                                                                                                PTPTT-----IKSAPTTPKEPAPTTT-------KSAPTTPKEPAPTTTKEP------A
                                         PDT--RPAPGSTAPPAHGVT-SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                            PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK
                                                                                                                                                                                                                                                                                                                                     1088
                                                                                                                                                                                                                                                                                                                                                                                  1163
1187
81
1034
957
975
1029
1055
11133
19
20
20
126
                                                                                                                                                                                                                                                                                                                                                             1077
                                                                                                                                                                                                                                                                      3 1255
2 2
1 134
1 154
1 1021
1 1021
1 1251
1 1261
1 1272 |
                                                                                                                                                                                                                                                                                                                                                            1087
                                                                                                                                                                                                                                                                                                                                                                                            1255
1162
1186
1186
1255
1255
1151
957
975
1029
1055
1133
1133
                                                                                                                                                                                                                              13.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALT_INIT.
ALT_INIT.
ALT_SEQ.
                                                                                                                                                                                                                    79;
                                                                                                                                                                                                                                                                   MISSING (IN SECRETED I

T -> A (IN REF. 11).

P -> Q (IN REF. 9).

P -> Q (IN REF. 9).

S -> T (IN REF. 3).

A -> T (IN REF. 3).

A -> T (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALT_SEQ
ALT_SEQ
ALT_SEQ
                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC...) (POTENI
T -> TATTAPKPAT (IN ISOFORM B)
MISSING (IN ISOFORM D).
MISSING (IN ISOFORM D).
MISSING (IN POLYMORPHIC EPITH)
                                                                                                                                                                                                                 Score 651; DB 1;
Pred. No. 3.4e-19
79; Mismatches 43
                                                                                                                                                                                                                                                                                                                                                 FLQIYKQGGFL -> VSIGLSFPMLP (IN ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EXTRACELLULAR (POTENTIAL)
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      44 X 20 AA TANDEM REPEATS
                                                                                                                                                                                                                  437;
                                                                                                                                                                                                                                      Length 1255;
                                                                                                                                                                                                                                                                      CRC64;
                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                    ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                    ) (POTENTIAL).
) (POTENTIAL).
) (POTENTIAL).
) (POTENTIAL).
) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                               EPITHELIAL
                                                                                                                                                                                                                                                                                                                                                                                                             В).
                                                                                                                                                                                                                 266;
                                                                                                                                                                                                                                                                                                                                                        SECRETED
                                                                                                                                                                                                               Gaps
                                                                                   181
                                                                                                      206
                                                                                                                             125
                     289
                                         238
                                                              252
                                                                                                                                                                      65
                                                                                                                                                                                          133
                                                                                                                                                                                                                52;
```

```
RESULT 5

FP1_MYTED STANDARD; PRT; 875 AA.

AC Q25460;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ADHESIVE PLAQUE MATRIX PROTEIN (POLYPHENOLIC ADH
DE PROTEIN 1) (MEFP1) (FRAGMENT).

GN MYTLIUS edulis (Blue mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomo
OC Mytlloidea; Mytilidae; Mytilus.

OX NCBL TaxID-6550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91025829; PubMed-1367451;
RR FILDUIA D.R. Lees S.M., Link R.P., Strausberg S.1

RT "Structural and functional repetition in a marine
                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                               Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                             QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ф
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ΩÝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ър
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                  991
                                                                                                                                                                                                                                                                                                                                                                                                                                   931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                290
                                                                                                                                                                                                                                                                                                                                                      PFSIPSHHSDTPTTLASHSTKTDASSTHHSSVPPLTSSNHSTSPQLSTGVSF
                                                                                                                                                                                                                                                                                                                                                                                    -----LVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDY
                                                                                                                                                                                                                                                                                                                                                                                                                      TAPPAHGYTSAPDNRPALGSTAPPVHNYTSASGSASGSASTLVHNGTSARATTTPASKST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITTT--EIMKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTP--FKITTLKTTTLAPKVTTTKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                               TPTPRKMTSTMPELNP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAP---GSTAPPAHGVTSAPDTRP---APGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAP----DTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --KKPAPTTPETPPPTTSEVSTPTTTKEPTT-----THKSPDESTPELSAEPTPKALE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NSPKEPGVPTTKTPAA-----TKPEMTTTAKDK---TTERDLRTTPETT-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDTRPAPGS--TAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DTRPAPGSTAP-PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -KKPAPKELAPTTKGPTSTTSDKPAP--TTPK----ETAPTTPKEPAPTTP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AP----PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAPDTRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AP--TTPKGTAPTTPKEPAPTTPKEPAPTTPKG----TAPTTLKEPAPTTP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPKETAPTTPKGTAPTTLKEPAPTTP-----KKPAPKELAPTTTKEPTSTTSDKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APGSTAPPAHGVTS----APDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---EPAPTTTKKPAPTAPKEPAPTTPKETAPTTPEKLAPTTPEKPAPTTPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LAPTTPEEPTPTTPEEPAPTTPKA----AAPNTPKEPAPTTPK----EPAPTTPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   T-SAPDT--RPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TKSAPTTPKEPSPTTTKEP-----APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KPAPTTPKEPAPTTPK----EPAPTTTKEPSPTTPKEPAPTT---TKSAPTTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----TSRIAEAMLQ--TTTRPNQTPNSK--
                                                                                                                     Pteriomorphia;
                                                                                                                                                                                       ADHESIVE
                   S
.L., Strausberg R.
ne mussel adhesive
                                                                                                                                                                                    PROTEIN) (FOOT
                                                                                                                   Mytiloida;
                                                                                                                                                                                                                                                                                                                                                                                        925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              733
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     448
```

functional repetition

Ø marine

밁

1167

VPTTYV

-TTLAPSAPVTPATNAVPTTITTTECSAATNAAGET

```
MUCCLLING MARKET TO SEE THE REAL MARKET TO SEE THE SEE THE REAL MARKET TO SEE THE SEE THE SECOND TO SEE THE REAL MARKET TO SEE THE SECOND TO SEE THE SECOND TO SECOND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MUCIN (TUMOR ASSOCIATED EPITHELIAL MEMBRANE ANTIGEN)

REACTIVE URINARY MUCIN) (PUM) (BREAST CARCINOMA-ASSOCIATED ANTIGEN

REACTIVE URINARY MUCIN) (PUM) (BREAST CARCINOMA-ASSOCIATED ANTIGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ungtenberg M.J.L., Vos H.L., Gennissen A.M.C., "Episialin, a carcinoma-associated mucin, is ge polymorphic gene encoding splice variants with termini.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-90368715; Pubmed-1697589;
Gendler S.J., Lancaster C.A., Taylor-Papadi
Peat N., Burchell J., Pemberton L., Lalani
"Molecular cloning and expression of human
polymorphic epithelial mucin.";
I Biol Chem 265.15286-15202/1000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE=Pancreas;

MEDLINE=90368716; PubMed=2394722;

Lan M.S., Batra S.K., Qi W.-N., Metzgar R.S.,

"Cloning and sequencing of a human pancreatic

J. Biol. Chem. 265:15294-15299(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=91097524; PubMed=2268309;
Lancaster C.A., Peat N., Duhig T., Wilson Taylor-Papadimitriou J., Gendler S.J.;
Taylor-Papadimitrion of the human pol gene: an expressed VNTR unit.";
Biochem. Biophys. Res. Commun. 173:1019-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Breast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                         Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J.,
Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.;
"A transcribed gene, containing a variable number of tandem repeats,
codes for a human epithelial tumor antigen. cDNA cloning, expression
of the transfected gene and over-expression in breast cancer
                                                                                                                                                                                                                                           "Human epithelial tumor antigen cDNA s
may generate multiple protein forms.";
Eur. J. Biochem. 189:463-473(1990).
                                                                                                                                                                                                                                                                                                                                           MEDLINE-90276413; PubMed-2351132;
Wreschener D.H., Hareuveni M., Tsarfaty
Horov J., Zaretsky J., Kotkes P., Weiss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM
                                                                                                                                           TISSUE=Breast carcinoma;
MEDLINE=90276414; PubMed=2112460;
                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                      Keydar
                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE=Breast carcinoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
       transfected
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Metazoa;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  carcinoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          265:5573-5578(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   265:15286-15293(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chordata;
Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Commun. 173:1019-1029(1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata; Vertebrata; Euteleostomi; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Taylor-Papadimitriou on L., Lalani E.-N., V
                                                                                                                                                                                                                                                                                                 sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polymorphic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            n, is generated b
                                                                                                                                                                                                                                                                                                                                                     ж.,
                                                                                                                              Kotkes P., Horev
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tumor-associated
                                                                                                                                                                                                                                                                                                                                                        Lathe
                                                                                                                                                                                                                                                                                                                                                                               Smorodinsky
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hollingsworth M.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alternative
                                                                                                                                                                                                                                                                                                       Differential
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wilson
                                                                                                                                                                                                                                                                                                                                                        R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ٦.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       epithelial mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mucin cDNA.";
                                                                                                                                                                                                                                                                                                                                                          Dion A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Duhig on D.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J.;
by a
                                                                                                                                                                                                                                                                                                                                                                                  z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T.,
                                                                                                                                                                                                                                                                                                         splicing
```

```
Gene
[8]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Tsarfaty I., Hareuveni M., Horev J., Z
Jeltsch J.M., Garnier J.M., Lathe R.,
"Isolation and characterization of an
coding for a breast-cancer-associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Abe M., Siddiqui J., Kufe D.;
"Sequence analysis of the 5' region
carcinoma-associated antigen gene."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=91033045; PubMed=1688329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         expressed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "A highly immunogenic region of a human polyn expressed by carcinomas is made up of tandem
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "preoperative diagnosis of thyroid
transcriptase polymerase chain reac
int. J. Cancer 66:55-59(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=96183746; PubMed=8608966; Weiss M., Baruch A., Keydar I., W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Thyroid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 1-109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biochem. Biophys. Res. Commun.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=90088473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1-169 FROM N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Burchell J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gendler S.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=88330762;
                                                                                                                                                                                                                                                                                                                                                                                                       Oncology 53:118-126(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                "Mucin mRNA expression in lung adenocarcinoma
                                                                                                                                                                                                                                                                                                                                                                                                                                             Yu C.J., Yang P.C., Shew J.Y., Lee L.N., Luh K.T., Wu C.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=96181716; PubMed=8604237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 J. Biol.
                                                                                                                                       This
                                                                                                                                                                                                                                                                                                                                 -
                                                                                                                                                                                                                                                                                                                                             Buluwela L., Liu Q.,
Submitted (OCT-1992)
                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-46 FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Lung;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE OF 1-89
 EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Breast carcinoma;
                                                                                   modified and this statement is not removed.
                                                                                                               the
                                                                                                                             between
                                                             or send an
                                                                         entities requires a
                                                                                                                                                              ACID).
POLYMORPHISM: THE NUMBE VARIES FROM 21 TO 125 I MOST FREQUENT ALLELES C MOST FREQUENT ALLELES C
                                                                                                                                                                                                                                                                                                                            uwela L., Liu Q., Luqmani Y.A., Gomm J.J., Coombes R. mitted (OCT-1992) to the EMBL/GenBank/DDBJ databases. FUNCTION: DIRECT OR INDIRECT INTERACTION WITH ACTIN
                                                                                                     s SWISS-PROT entry is copyright. It is produced through a collab ween the Swiss Institute of Bioinformatics and the EMBL outst Buropean Bioinformatics Institute. There are no restrictions by non-profit institutions as long as its content is in
                                                                                                                                                                                                                           TISSUE SPECIFICITY: ABERRANTLY EXPRESSED TUMORS, SUCH AS BREAST CANCER.
PTM: HIGHLY GLYCOSYLATED (N-AND O-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J. Biochem. 189:475-486(1990)
                                                                                                                                                                                                                                                                   SPLICING
                                                                                                                                                                                                                                                                               ALTERNATIVE
                                                                                                                                                                                                                                                                                                          SUBCELLULAR
                                                                                                                                                                                                                                                                                                                     CYTOSKELETON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93:313-318(1990).
                                                                                                                                                                                                                                                                                            IS ALSO PRODUCED
 J05582;
M32738;
M32739;
J05581;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM
                                                                                                  non-profit institutions as long
                                                             equires a license agreement (email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCE FROM N.A.
762; PubMed=3417635;
Taylor-Papadimitriou
                          AAA60019.1;
AAA35804.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 rcinomas is made up of 263:12820-12823(1988).
  AAA35806.1;
AAA59876.1;
                                                                                                                                                                                                                                                                                PRODUCTS: VARIOUS
                                                                                                                                                                                                                                                                                                        LOCATION: TYPE I MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PubMed=2597151;
                                                                                                                                                                           THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC. 1 TO 125 IN THE NORTHERN EUROPEAN POPULATION. ALLELES CONTAINS 41 AND 85 REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                 N.A.
                                                                                                                                                                  1 SEA DOMAIN
                                                                                                                                                                                                                              (N-AND O-LINKED CARBOHYDRATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hong
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     roid papillary carcinoma by reaction of the MUC1 gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            165:644-649(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wreschner D.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  J.,
                                                                                                                                                                                                                                                                                  VARIANTS ARE PRODUCED BY ALTERNATIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Т.М.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zaretsky J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               expressed hypervariable antigen.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Keydar I.,
                                                                              (See
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Duhig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           polymorphic epithelial mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yang
                                                                                                                                                                                                                                                                                                           PROTEIN.
                                                                              http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               repeats.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   T., Rothbard
                                                                                           Usage
                                                                                                                                                                                                                                                           IN HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                               S.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wreschner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DF3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weiss
                                                                                                                                                                                                                                                                                                              Þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lee
                                                                                                                                                                                                                                                                                                              SECRETED
                                                                                                                                                                                                                                                            EPITHELIAL
                                                                                           and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Y.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     reverse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            D.H.;
                                                                                                                                               a collaboration
                                                                                              for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gene
                                                                                                                                                                                                                                                                                                                FORM
                                                                                                                                  outstation
                                                                                                                                                                                                BHT
                                                                                              n no way
```

S

```
ARMINITERS 3

ANTINITERS 31995 (Rel. 09, Created)

DT 0-FEB 1995 (Rel. 31, Last sequence update)

DT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Barrell B.G., Badcock K., Bankler A.T., Bowman S., Brown D., Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D., Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V., Walsh S.V., Whitehead S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Saccharomyces cerevisiae (Baker's yeast).

Bukaryota; Fungl; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

NCBI_TaxID-4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S SWISS-PROT entry is copyright. It is produced through a collab ween the Swiss Institute of Bioinformatics and the EMBL outst European Bioinformatics Institute. There are no restrictions by non-profit institutions as long as its content is in the content of th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------KPTSTPAPTEIE-EPTPSDVPGAIGGEHRAYLRGY------PDGSFRP 1465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
GLUCOAMYLASE S1/S2.
SER/THR-RICH.
N-LINKED (GLCNAC. .
        (POTENTIAL)
                                                                                                                                                                                                      Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         evolution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          collaboration -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              X >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0n
```

```
B
                                                          δÃ
                                                                                                              ρ
                                                                                                                                            Вþ
                                                                                                                                                                      δÃ
                                                                                                                                                                                                    Ъ
                                                                                                                                                                                                                             ρ
                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                           δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
Matches 280
                                  1107
                                                                                        1047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
SEQUENCE
     864
                                                                                                                 769
                                                                                                                                              987
                                                                                                                                                                                                    932
                                                                                                                                                                                                                                                            875
                                                                                                                                                                                                                                                                                       606
                                                                                                                                                                                                                                                                                                                                                                          758
                                                                                                                                                                                                                                                                                                                                                                                                      494
                                                                                                                                                                                                                                                                                                                                                                                                                                   869
                                                                                                                                                                                                                                                                                                                                                                                                                                                              447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      296 TKKKTTTSK-TCT---KKTTTPVPTPS-SSTTESSSAPV-----PTPSSSTTESSSAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   35
                       SGCSPKTVTTTVPCSTGTGEYTTEATTLVTTAVTTTVVTTESSTGTNSAGKTTTGYTTKS
NPTSRIAEAMLQTTTRPNQ--TPNSKLVEVNPKSED
                                                  --AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPEL
                                                                             STSTKPGGEITTTFVTKNIPTTYLTTIAPTPSVTTVTNFTPTTITTTVCSTGTNSAGETT
                                                                                                                                  VCSTGTNSAGETTSGCSPKTITTTVPCSTSPSETASESTTTSPTTPVTTVVSTTVVTTEY 1046
                                                                                                                                                           ---TAKDKTTERDLRTTPET-TTAAPKMTKETATTTEKTTESKTTATTTQVTST--TTQD
                                                                                                                                                                                            TTPSTTTTTTVCSTGTNSAGETTSGCSPKTVTTT----
                                                                                                                                                                                                                      STPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTT----
                                                                                                                                                                                                                                                                           TTTKGP----TSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPET---PPPTTSEV
                                                                                                                 TTPFK---
                                                                                                                                                                                                                                                  ITSSAPSSIPFSSTTESFSTGTT----VTPSSSKYPGSQTETSVSSTTETTIVPTKTTTSV
                                                                                                                                                                                                                                                                                                         ITSSAPSSTPFSSSTESSSVPVPT----PSSSTTESSSAPVSSSTTESSVAPVPTPSSSSN
                                                                                                                                                                                                                                                                                                                                 TSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --PSSSTTESSSAPVPTPSSSTTESSSAP-VPTPSSSTTESSSAPVTSSTTESSSAPVTS
                                                                                                                                                                                                                                                                                                                                                                 ESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSVAPVPTPSSSSN
                                                                                                                                                                                                                                                                                                                                                                                                                        STTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVTSSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EELAPTTPEEPTP---TTPEEPAPT-TPKAAAPNTPKEPAP----TTPKEPAPTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAP----TTPEKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -PSSSTTESSSAPVPTPSSSTTESSSAPAPTPSSSTTESSSAPVTSSTTESSSAPVPT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTT-----PKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTAPKKPAPTTPKEPAPTTPKE-----PAPTTTKEPSPTTPKEPAPTTTKSAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEPAP----TTTKEPAP----TTPKEPAP----TTTKEPAP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APAPTPSSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSSTPVTSSTTESSSAPVPT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TTESSSAPVTSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SSSAPVTSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LTTPKE-----PTPTTPKEPASTTPKEPTPTTIKSAP----TTPKEPAPTTTKSAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTPKKPAPTTPKEPAPTTPKEPTP---TTPKEPAPTTKE-----PAPT----TPKEPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSSTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STSESSTTTSSTSESSTSSSTTAPATPTTTSCTKEKPTPPTTTSCTKEKPTPPHHDTTPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETT--TKGPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTPDTSTTQHNKVSTSPKITTAKPINP------RPSLPPNSDTSKETSLTVNKETTVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                             -PKEPAPTTPKETAPT-TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 874
1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ΑA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            874
136110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         15.2%;
27.8%;
                                                                                                       -ITTLKTTTLA--PKVTT----TKKTITTTEI---MNKPEET-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC. . . . MW; 91C00E2DBD61AA9D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 762.5; DI
Pred. No. 1.7e-1
15; Mismatches (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1;
                                                                                                                                                                                            - VPTTTTTSVTTSSTTTITTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .) (POTENTIAL)
) CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -TTTKSAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                            986
                                                                                                                                                                                                                                                   931
                                                                                                                                                                                                                                                                              658
                                                                                                                                                                                                                                                                                                                                      605
                                                                                                                                                                                                                                                                                                                                                                                             545
                                                                                                                                                                                                                                                                                                           874
                                                                                                                                                                                                                                                                                                                                                                 817
                                                                                                                                                                                                                                                                                                                                                                                                                                                     493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               697
                                                                                                                                                                                                                                                                                                                                                                                                                        757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42;
```

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         БР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qγ
STEPLICE
STE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SLP1_CLOTM
Q06852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2244
                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                     STRAINNCE 10682;

MEDLINE-9330931; PubMed-8458832;

Fujino T., Beguin P., Aubert J.-P.;

Fujino T., Beguin P., Aubert J.-P.;

Toganization of a Clostridium thermocellum gene cluster encoding cellulosomal scaffolding protein CipA and a protein possibly involution attachment of the cellulosome to the cell surface.";

J. Bacteriol. 175:1891-1899(1993).

J. Bacteriol. 175:1891-1899(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-1994 (Rel. 29, Created)
01-JUN-1994 (Rel. 29, Last sequence up
01-FEB-1996 (Rel. 33, Last annotation
CELL SURFACE GLYCOPROTEIN 1 PRECURSOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   551
  Pfam; PF(PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bacteria; Firmicutes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEIN 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=1515,
                           InterPro; IPR001119; Spfam; PF00395; SLH; 3.
                                                                              EMBL; X67506; CAA47841.1;
                                                                                                                                                                                                                                                                                                                                          -!- SIMILARITY: CONTAINS 4 S-LAYER HOMOLOGY (SLH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 lostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTPKEPAPT-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TPTTTPITTTTVTPTPTPTGTQTPT-TTPITTTTTVTPTPTPTGTQTPTTTTTTTTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AP-TTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTK
                                                                                                                                                                                                                                                                                                                                                                        SUBCELLULAR LOCATION: CELL WALL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIHKSPDESTPELSAEPTPKALENSPKEP-----GVPT-TKTPAATKPEMTTTAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GPTSTTSDKPAP-TTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTTVTPTPTGTQTPTTTPTTTTTTVTPTPTGTQTPTTTPITTTTTTTTTPTPTFTGTQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TPFKITTLKTTTLAPKVTTT-KKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DKTTERDLR---TTP--ETTTAAPKMT---KETATTTEKTTESKITATTTQVTSTTTQDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-KKPTSTKKDKTMPRVRKPKTTPTPRKMTSTMPELNP---TSRIAEAMLQTTTR-PNQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PITTTTTVTPTPTG----TOTP--TTTPITTTTTVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTP 921
        PS01072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     thermocellum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
     SLH_DOMAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacillus/Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              update)
(OUTER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Clostridiaceae,
                                                                                                                                                                                                                                                                                                                                                     DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              coding the involved
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (S-LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 609
```

```
δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                              Db
                                                                                                                                                                                                                                  δδ
                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                     δã
                                                                                                                                                                                                                                                                                                                                          Вb
                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                               δÃ
                                                                                                                                                                    B
                                                                                                                                                                                         δõ
                                                                                                                                                                                                                                                        밁
                                                         δÃ
                                                                                  밁
                                                                                                     γ
                                                                                                                            밁
맑
                  δÃ
                                        뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 276; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                1123
                                                                                                                                                                                                                                                          1065
                                                                                                                                                                                                                                                                                                    1020
1373
                                                                                   1278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     128 VLAKPTP-KAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTTKSAPTTPKEPAPTT
                                                                                                                             1230
                                                                                                                                                                                                                                                                                                                                                                                     920
                                                                                                                                                                                                                                                                                                                                                                                                                              865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            187
                                                                                                                                                                                                                                                                                                                       417
                                                                                                                                                                                                                                                                                                                                                                                                         302
                                                                                                                                                                                           592
                                                                                                                                                                                                                                     533
                                                                                                                                                                                                                                                                               476
                                                                                                                                                                                                                                                                                                                                            969
                                                                                                                                                                                                                                                                                                                                                                358
                                                              767
                                                                                                       707
                                                                                                                                                 647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TKSAPTTPKEPAPTTTKEPAPTTPKEPA----PTTTKEPAPTTTKSAPTTPKEPAPTTPK 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VVIQPAPIKAASDEPIPTDTPSDEPTPS-
                                                                                                                                                                                                                                                                                                                                                                                               SDEPTPSDEPTPS--DEPTP-SDEPTPSETPEEPIPTDTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                         KPAPTTPKEPAPTTPKEPAPTTPKEPAPTTEPAPTTPKEPAPTAPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EPTPSETPEEPIPTOTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTOTP
                                                                                                                                                                                                                                                                                                                                          SDEPTPS---DEPTPS----DEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSD---
                                                                                                                                                                                                                                                                                                                                                              TKEPAPTTPKEPAPTTPKKPAPTTPKEPAPT-TPKEPAPTTTKKPAPTAPKEPAPTTPKE
                                                                                                                            DTPSDEPTPSD--EPTPSDEPT----PSDEPTP--SDEPTP--
                                                                                                                                        PT-TPKKPAPKELAPTTTKGPTSTTSDKPAPT---TPKETAPTTPKEPAPT-TPKKPAPT | : | : | : | : | : | : | : | : | | : | : | | : |
                                                                                                                                                                                                               D-EPTPSDEPTPSETPEEPIPTDTPSDEPTPS--DEPTPS---DEPTPS----DEPTPSDEPT
                                                                                                                                                                                                                                  ELAPTTTKEPT-STTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA
                                                                                                                                                                                                                                                         TPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPS--DEPTPS
                                                                                                                                                                                                                                                                           NTPKEPAPTTPKEPA--PTTPKEPAPT-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPK
                                                                                                                                                                                                                                                                                                                     TAPTTPKKLIPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPT-TPEEPAPTTPKAAAP
                                                                                                                                                                                                                                                                                                                                                                                     SDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPS---
                                                                                  EPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPS
                                                                                                      ATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTT
SETPEEPTPTTTP---
                                                            QDTTPFKITTLKTTTLAPKVTTTKKTITTEIMNKPEETAKPKDRATNSKATTPKPQ-KP
                    TKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S-la
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1378
1453
1495
1566
1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29
36
36
207
409
607
771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                       ----IPTDTPSDEPTPSDEPTPSDEPTPSD-EPTPSDEPTPSETPEEPIPT
                                                                                                                                                                                                                                                                                                  EPTPSDE---PTPSDEPTPSD----EPTPSDEPTPSETPEEPTPTDTPSDEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            yer;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1449
1494
1565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28
1664
763
191
363
565
763
1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  178194
 -TPTPSTTPTSGSGGSGGSGGGGGGGTVPTSPTPTPTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPROXIMATE TANDEM REI
T-P-S-D-E-P
GLK/PRO/SER/THR-RICH.
SLH 1 (INCOMPLETE)
SLH 2
SLH 3
SLH 4 (INCOMPLETE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Repeat.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 942; DB 1;
Pred. No. 2.2e-30;
3; Mismatches 312;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ELL SURFACE GLYCOPROTEIN 1.
x 156 AA APPROXIMATE REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4 (INCOMPLETE).
5F396695BA9FE74B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---DEPTPS---DEPTPSDEPTPSD 804
                                             PSDEPTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1664;
                                                                                                                                 -SETPEEP-IPTDTPSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                        TMPELNPTSRIA
                                                                                                                                                                                                                                                                                                                                                                                         ----DEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                     1175
                                                                                                                                                                                                                                                                                    532
                                                                                                                                                                                                                                                                                                                             475
                                                                                                                                                                                                                                                                                                                                                 1019
                                                                                                                                                                                                 646
                                                                                                                                                                                                                                                                1122
                                                                                                                                                                                                                                                                                                         1064
                                                                                                                                                                                                                                                                                                                                                                       416
                                                                                                                                                                                                                                                                                                                                                                                          968
                                                                                                                                                                                                                                                                                                                                                                                                                                     919
                                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            186
                                                                                                                                                                                                                                          591
                                                                                                                                                                                                                                                                                                                                                                                                                357
     1424
                                               1372
                                                                                         1337
                                                                                                             766
                                                                                                                                    1277
                                                                                                                                                         706
                                                                                                                                                                            1229
                           870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  43;
```

```
PRINTS; PRO0438; GFCYSKNOT.
SMARR; SM00214; VWC; 2.
SMARR; SM000214; VWC; 4.
SMART; SM000215; VWD; 4.
PROSITE; PS00022; EGF_1; UNKNOWN_1.
PROSITE; PS01125; CTCK_2; 1.
PROSITE; PS01126; VWFC; 2.
                                           REPEAT
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
CARBOHYD
CARBO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000359; Cys_knot.
InterPro; IPR000561; EGF-11ke.
InterPro; IPR002400; GF Gysknot.
InterPro; IPR001007; VWFC.
InterPro; IPR001846; Vwd.
Pfam; PF00007; Cys_knot; 1.
Pfam; PF000094; Vwd; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L21998
EMBL; M74027
EMBL; M94131
EMBL; M94132
MIM; 158370;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
DOMAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glycoprotein;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L21998; AAB95295.1; -. M74027; AAA59875.1; -. M94131; AAA59163.1; -. M94132; AAA59164.1; -.
                                                                                                163
423
670
770
794
1139
1154
1215
1230
                                                                                                                                                                                                                                                                                                                                                                                    Repeat;
                                           1416
14432
14432
14432
14474
15556
15556
15596
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
16635
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20
5179
1747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Signal
                             POTENTIAL.

MUCIN 2.

APPROXIMATE REPEATS.
1.
2.
3.
4.
4.
4.
4.
4.
5.
6.
6.
7 A.
7 B.
8 B.
9 B.
9 B.
11 A.
11 B.
11 A.
11 B.
11 A.
11 B.
11 B.
11 A.
11 B.
11 B.
11 A.
11 B.
11 A.
11 B.
11 B.
11 A.
11 A.
11 B.
11 A.
11 A.
11 B.
11 A.
1
                                                    (POTENTIAL)
                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                       망
                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local
1890 TENPTEPTTTPITTTTTTTPTPTPTGTQTPTTTPITTTTTTTTPTPTPTPTGTQTPTTTPITT
                                                                                                                                                                                                                                             1770
                                                                                                                                                                                                                                                                                                                                                       1710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CONFLICT
CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1312
                                                                                                                                                                                                                                                                                               415
                                                                                                                                                                                                                                                                                                                                                                                                              356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PEKPAPTTPEELAPTTPEEPTPT-TPEEPAPTTPKAAAANTPKEPAPTTPKEP----AP
                                                                                                         DVPIGQLGQTVVCDVSVGLICKNEDQKPGGVIPMAFCLNYEINVQCCECVTQPTTMTTTT
                                                                                                                                                                                                                      FSPFSTTTPTTPCVPLCNWTGWLDSGKPNFHKPGGDTELIGDVCGPGWAANISCRATMYP
                                                                                                                                                                                                                                                                                                                               ----KETAPTTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KEPAPTTTKEPAPTTPKEP-----APTTTKEPAPTTTKSAP--TTPKEPAPTTPKKPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEPT-----PTTPKEPASTTPKEPTPTTIKSAP-TTPKEP-----APTTTKSAP-TTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ITTTTPPPTTTTSPPTTTTTTPPPTTTPSPPTTTPIPTSTTTLPPTTTTPSPPPTTTTT 1650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEP------APTAPKKPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNGPFGLCYDYKIRVNCCWPMDKCITTPSPPTTTPSPPPTTTTLPPTTTPSPPTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DHPSSGSDDGDREPFDGVCGAPEDIECRSVKDPHLSLEQHGQKVQCDVSVGFICKNEDQF 1371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ETTTTNKQTSTDGKEKTTSAKETQS1EKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4339
43621
43623
44373
44438
4438
4438
44502
44502
44502
4627
4752
4752
4888
4975
4985
5038
5038
5038
5038
5038
51351
1412
11412
11412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32
23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                540295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 1168; D
Pred. No. 9.1e
53; Mismatches
                                                                                                                                                                                                                                                                             -----KKLTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; DB 1;
9.1e-39;
nes 448;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 5179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                    -TTPEKLAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                          491
                                                                                                            1889
                                                                                                                                                                                                                                                                             427
                                                                                                                                                                                                                           1829
                                                                                                                                                                                                                                                                                                                                     1769
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40
```

```
OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
                                                                                                                                                                    Score
  442.5
439.5
432.5
432.5
432
424.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ı
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-556-246-1_COPY_200_1140
5011
1 VKDNKKNRTKKKPTPKPPVV........DMDYLPRVPNQGIIINPMLS 941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 April 26, 2002, 16:13:35; Search time 47.4 Seconds (without alignments) 727.882 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100059 seqs, 36664827 residues
    100.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
110.66
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SwissProt_39:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd
    DВ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gapext 0.5
RPBI_HUMAN
RPBI_MOUSE
NFH_RAT
RPBI_CRIGR
SSP2_PLAYO
NFH_HUMAIZE
RYTN_MAIZE
HWPI_CANAL
EXTN_MAIZE
VRPI_YEAST
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
YJJYAN
YJ9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                              EXTN_TOBAC
NFH_MOUSE
TCNA_TRYCR
CPN_DROME
FP1_MYTCO
MUC1_XENLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100059
                                                                                                                                                                        Q25434 mytilus cor Q25434 mytilus cor Q25049 xenopus lae p24928 homo sapien p08775 mus musculu p16884 rattus norv p11414 cricetulus coll443 plasmodium p12036 homo sapien p14018 zea mays (m p46593 candida alb 088799 mus musculu p37370 saccharomyc p487179 saccharomyc p48719 homo sapien Q00130 ictalurid h
                                                                                                                                                                                                                                                                                                                                                                                                                               P19246
P23253
P232510
Q02910
Q25434
Q05049
P24928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q25460
Q9Y493
P13983
           P37370
P47179
P48634
Q00130
P28968
Q27409
000268
P17437
P10220
Q9umn6
P27546
P02840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0 saccharomyc
1 h mucin 1 p
0 mytilus edu
3 homo sapien
3 nicotiana t
           gequine herp
mytilus gal
homo sapien
renopus lae
herpes simp
homo sapien
mus musculu
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 homo sapien
clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             trypanosoma
drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nicotiana t
mus musculu
```

000000000000000000000000000000000000000	~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	MODE MODE MODE MODE MODE MODE MODE MODE	
			335 344 300 300 300 300 300 300 300 300 300
J. CLIH. LIVES	SEQUENCE OF 626-18 TISSUE=COlon; MEDLINE=03016075; Gum J.R. Jr. Hick Kin Y.S.; The human MUC2 in both upstream and J. Biol. Chem. 267 [3] MEDLINE=91358717; Toribara N.W., Gum petersen G.M., Kin "MUC-2 human small and polymorphism."	MUC2-HUMAN MUC2-HUMAN MUC2-HUMAN O1-NOV-1994 (Rel. 29, Created) O1-NOV-1997 (Rel. 35, Lasted) 20-AUG-2001 (Rel. 40, Lasted) MUC1N 2 PRECURSOR (INTESTILM) HOMO Sapiens (Human). Eukaryota: Metazoa; Chorda Mammalia; Eutheria; Primated NCBL-TaxID-9606; [1] TISSUE-Intestine; MEDLINE-94132002; PubMed-8 MEDLINE-94132002; PubMed-8 Gum J.R. Jr., Hicks J.W., "Molecular cloning of human Identification of the amin to prepro-von Willebrand for Bioli. Chem. 269:2440-24	422.5 422.5 419.5 419.4 11.5 411.6 410.6 406 406.6 400.5
FUNCTION: COATS THE PUNCTION: COATS THE POOTER MUCUS MEMBER POOTECTIVE, LUBRIG AGENTS AT MUCOSAL SUBJUNT: MULTIMERI SUBCELLULAR LOCATI TISSUE SPECIFICITY SISSUE SPECIFICITY TISSUE SPECIFICITY SPECIFICATI ALL CYSTEINE PROLYMORPHISM: THE POLYMORPHISM: THE VARIES AMONG DIFFISMILARITY: THE NOF SIMILARITY: CONTA SIMILARITY: CONTA SIMILARITY: CONTA SIMILARITY: CONTA	OF 626-1899 clon: 75; Pt 93016075; Pt 93016075; Pt 93016075; Pt 10. Hicks an MGC2 inte tream and de tream and de tream and de tream 367: Chem. 267: OF 1343-188 OF	Q14878; 1994 (Rel. 2) 2001 (Rel. 2) 2001 (Rel. 4) PRECURSOR (Human ta; Metazoa; a; Eutheria, xID=9006; axID=9006; by 1015 (Human ta; Metazoa;	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
TITH. LINESC. SOLIDATIONS OF STRUCTION: COATS THE EPITHE PRINE PROTECTIVE, LUBRICATING BAPAREANE-CONTA PROTECTIVE, LUBRICATING BARDENIS AT MUCOSAL SURFACES SUBBUNIT: MULTIMERIC. SUBCELLULAR LOCATION: SECR SIBSUME SPECIFICIT: COLON, SIBCONCHUS, CERVIX AND GALLES PTM: ALL CYSTEINE RESIDUES PTM: ALL CYSTEINE RESIDUES PTM: ALL CYSTEINE RESIDUES PROLYMORPHISM: THE NUMBER CYARIES AMONG DIFFERENT ALL SIMILARITY: THE N-TERMINAL SIMILARITY: CONTAINS 2 WHESIMILARITY: CONTAINS 1 C-T	1:1895 AND 4 15; pubMed=1 1icks J.W., 2: intestinal 2: ind downstre 2: 1375-2 13-1895 AND 17; PubMed=1 16 Gum J.R. Jr 6 Gum J.R. Jr 6 J.R. Jr 8 Kim Y.S.; 8 Jr 8 J.R.	ANDARD; 29, Creat 35, Last 40, Last (INTESTII an), chorda a; Chorda a; Primat PubMed=8 PubMed=8 Pks J W., syg of huma the amin the amin tlebrand f	1251 1794 3476 1185 1185 1229 929 1229 1229 133 2774 1183 2774 143 2774 143 2774
TING TING TING TING TING CO. CO. CO. CO. TING TING TING TING TING TING TING TING	ND 419 W., To W., To inal stream stream 75-213 AND 41 led=188 led=188 led=188); Preated; Preated; STINAL ST	444444444444444444444444444444444444444
LIA OF THE INTESTINE LIA OF THE INTESTINE INING ORGANS. THOUGH RRIER AGAINST PARTIC SMALL INTESTINE, CO SMALL INTESTINE, CO SMALL INTESTINE, CO BLADDER. ARE INVOLVED IN INT (BY SIMILARITY). (C DOMAIN SHOWS SOME S C DOMAINS.	OM N.A. W., Rothe EM. cysteine-rich s entral repetiti ROM N.A. P.J., Lagace R. pene structure.	PRT; 5179 AA. ad) sequence update) annotation update) AL MUCIN 2). 3; Craniata; Vertebrata; 5; Catarrhini; Hominidae; 60571; Onibara N.W., Siddiki B., intestinal mucin (MUC2) terminus and overall secotor."; 6(1994).	YQU3_CAEEL YAV1_SCHPO TEGU_HSVEB ZAN_PIG DRPL_HUMAN N1_21_HUMAN VGP3_EBV URPL_RAT MAPA_RAT K167_HUMAN TCOF_HUMAN NCR2_HUMAN NCR2_HUMAN ALIGNMENTS (
NES, AIRWAYS, AND GHT TO PROVIDE A ICLES AND INFECTIOUS COLONIC TUMORS, NTRACHAIN OR AY POLYMORPHIC AND SIMILARITY TO THAT SIMILARITY TO THAT	., Lagace R.E., subdomains located ive region.";	e; Homo. e; Homo. color Y.S.;	009550 caenorhabdi 010172 schizosacch p28955 equine herp 028983 sus scrofa p54259 homo sapien 0992n3 homo sapien 0992n3 homo sapien p03200 epstein bar p54258 rattus norv p34926 rattus norv p46013 homo sapien 013428 homo sapien 013428 homo sapien

```
밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          á
                                                                                       878
                                                                                                                  811 RATINS---KATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTP-RKMTSTMPELNPT
                                                                                                                                                                836 KNTSTTKIANST-------PKAAAAECSEESSSEDEGKANGTSGKRKR
                                                                                                                                                                                                                                                                                            694 PKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESK 753
                                                                                                                                                                                                                                                                                                                                                   723
                                                                                                                                                                                                                                                                                                                                                                          636 APTTPKK-PAPTTPETPPPTTSEVSTPTTTKE-PTTTHKSPDESTPELSAEPTPKALENS
                                                                                                                                                                                                                                                                                                                                                                                                                             666 SEKK---TKPAKPPAKSATPVNTKAPAQNKASKASCSDSDSSSEEEGKSKQPTGKSPAAK 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      586 TLKEPAPTTPKKPAPKELAPTTTKGP-----TSTTSDKPAPTTPKETAPT--TPKEP 635
                                         867 SRIAE 871
                                                                                                                                                                                                       754 ITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTK--KTITTTEIMNKPEETA-KPKD 810
                                                                                                                                                                                                                                                              783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  612 KPVPAK-----KASSSSDSDSSEEETTKTTKPLTKLSPAVKTLPPKKAESSSDSSSDSD 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     559 AAKSTPGKQVPTKKESSSSDSSDSSSEDEKKSSAKPAVKTTPGKATS------KPVVAS 611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            515 SDSSDSSSDEKKTPA--KRAAKTTPAKPA-----AKTTPAKPA-----AKTTPAKP 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      401 TKPAAKTTPAKSAATPTSKTPINGK---ATPTSKTPAKPGTPKTSTAKKDSSSSDSSDSS 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  273 KEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTP----KEPAP------TT 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              284 LSQPGTKAKPESSDSSDESEDGPAKKAKIVPAKAASAPKPLAKKAETSTDSESDSSS 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   226 --KPATAAKPQAKKTAGKKSSSREDSSDSSDEEQKTAKSKPKPDVYSAVPPPTSVSKKKT 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173 KSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTT-TKEPAPTT----T 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       227 KSAPTTPKEPAPT-----TPKKPAPTTPKEPAPTTP----KEPTPTTPKEPAPTT 272
                                                                           ESTGNAECEAVTPE-----NKKLKAKSPNTFPKVNKKELKNTPFRRVVEEDIEINP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAPNTPKEPAPT----TPKEPAPTTPKEPAPTTPKETAPTTP-KGTAPTTLKEPAPTTPK 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----PTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTP-KA 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDEETTTKPAAKTTPAKSAATPTSKTP---TNSKATPTSKKTPAKPGTPKTSAAKKDSSS 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---PTTPKKPAPTTPKEPAPT-TPKEPAPTTTKKPAPTAPKEPA-PTTPKETA----- 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA- 370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EDEKKSSVKLGVKAAPKK-APAAP--DAKSTPVAAAKKSAPAKKASSSSDSDSSSNEETT 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKE----PASTTPKEPTPTTI 172
-RMAD 932
                                                                                                                                                                                                                                                   SSE----DSSSDEDVSKAKKTNTAVSKSPV----TTPKAVPAAKKESSSESSDSEDEKQGG
                                                                                                                                                                                                                                                                                                                                          ATAPPKKNPVAVNKDKPSSSSSSDSSGDDEKQKPKQAAAAKDVKQGAKAAKPTPKKAASS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KPAPKELAPTITKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT--PKEPAPTTPKGTAPT 585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EASKKKOPVIKV----PPKOAVVKAGLASNNGKTADSSSSEDSDSPPAKKTAATKTPPT- 225
                                                                                                                       866
                                                                                                                                                                                                                                                     835
                                                                                                                                                                                                                                                                                                                                          782
                                                                                                                                                                                                                                                                                                                                                                                      693
```

Search completed: April 26, 2002, 16:18:10 Job time: 280 sec

õ

밁

õ

В Ş 밁 Š Ş 밁 Ş Š

428 TTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPN 476	42	Qy
350 NLSQSSTVSTEDRSEISSENSEKPT-SAPELVTSSVTHVASSSPDVPTESSEPDDLTG 2406	235	Ф
98EPAPTTPKETAL	39	Qy
345 SAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTT- 397	345	dp Qy
APKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTK	287	DP 03
229 APTTPKEPAPTTPKKPAPTTPK-EPAPTTPKEPTPTTPKEPAPTTKEP-APTTPKEPAPT 286	229	Дb
170 TIIKS-APTIPKEPAPTTIKSAPITPKEPAPTITKEPAPTITKEPAPTITKS 228 	170 2120	da Vo
114 IEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTP 169 11 ::	114 2061	ОУ
58 PINPRPSLPPNSDTSKETSLTVNKETTV-ETKETTTTNKQTSTDGKEKTTSAKETQS 113 1	20	Оγ
10 KKKPTPKPPVVDEAGSGLDN-GDFKVTTPDTSTTQHNKVSTSPKITTAK 57 : :	19	Фр
y Match 12.6%; Score 629; DB 2; Length 3507; Local Similarity 24.1%; Pred. No. 1.7e-19; hes 252; Conservative 158; Mismatches 416; Indels 218; Gaps 45;	Query Ma Best Loc Matches	
ics: CESP:ZK783.1 osition: 3 ns: 14/1; 48/2; 84/1; 196/3; 303/1; 381/1; 5	Genetics Gene: CE Gene: CE Map posi Introns:	3 × × × C
Status: preliminary; translated from GB/EMBL/DDBJ Molecule type: DNA Molecule type: DNA Residues: 1-3507 <fav> Cross-references: EMBL:U13646; PIDN:AAC24418.1; GSPDB:GN00021; CESP:ZK783.1 Experimental source: strain Bristol N2; clone ZK783</fav>	Status: Molecule Residues Cross-re Experime	
R;Favello, A.; Vaudin, M. submitted to the EMBL Data Library, August 1994 s;Description: The sequence of C. elegans cosmid ZK783. A;Reference number: Z21536 A;Recession: T34513	Favello, A bmitted to Descriptio Reference Accession:	A A Su R
SSULT 14 94513 Ppothetical protein ZK783.1 - Caenorhabditis elegans Species: Caenorhabditis elegans pate: 29-Oct-1999	RESULT 14 T34513 hypothetical C;Species: Ca C;Date: 29-Oc C;Accession:	C;
)78 TRPNQTPNSKLVEVNPKSEDAGGAEGET 905 : ; : :	878 6973	дь Оу
127 KAPKKPTSTKKPKTMPRVRKPKT-TPTPRKMTSTMPELNPTSRIAEAMLQTT 877 	827 6915	DP GA
67 QDTTPFKITTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPT 826	767 6877	Оy

DЬ

B Qy

115

PSLPPNSDTSKETSLTVNKETTVE--TKETTTTNK----QTSTDGKEKTTSAKETQSIEK 116

```
멍
                                                                                                                                                                                                                                                                                                                                                           R; Cairns, C.; McStay, B.
J. Cell Sci. 108, 3339-3347, 1995
A; Title: Identification and cDNA cloning of a Xenopus
A; Reference number: IS1618; MUID:96019267
A; A; A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nucleolar phosphoprotein - African clawed frog C;Species: Xenopus laevis (African clawed frog) C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 C;Accession: I51618; S57757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Вρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Вþ
                                                                                                                                                                                                     A;Gene: xNopp180
C;Superfamily: nucleolus-cytoplasm shuttle phosphoprotein
C;Keywords: phosphoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вb
                                                                                                                                                                                                                                                                         A;Cross-references: EMBL:X88927; NID:g895920; PIDN:CAA61368.1; PID:g895921
C;Genetics:
                                       Qy
                                                                                                                                                                                                                                                                                                                    A; Molecule type: mRNA
A; Residues: 1-990 <CAI>
                                                                                         Query Match
Best Local Sin
Matches 253;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2465 --KSPVSSSTEGISVVTSTEFSKVPESTISSVLEEDLTKTTPSPILEETTTASETSEPLT 2522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2753 KLSSLFPESITSEAVTVSSR----APAEITMSSESHREISTVSSEPSEPEIPLSTTVSPN 2808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2703 SGSLESSTMSSTS-----SEPETNAPAVTVSSEASSTTLEENSSTSSP----TSSEASV 2752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2866 SNSTSPIILPSES-LTTPQPPPTTTTTAKPATTSGKRGPPSIQPPAEMFTTPAP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          860 MPELN-----PTSRIAEAMLQTTT 878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15
  59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTTSEVSTPTTTKEPTTIHKSPDESTP--ELSAEPTPKALE--NSPKEPGVPTTKTPAAT 708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEIEAITSNTPFKQGRTPITTSPKSLVKSTTSPSTVTSSEPSESTKRTTVSTTVSTTTPT 2642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TPKEPAPTT - - - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STENIPEASSKQTISSTPTPDTTTASEEPTKSTSMSPDLSTTSNVLSESSTTPESSS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PPPSNGGYGEETNQEEEQVTSTTT 2942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               K-----PEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EETTTSESLILTAAPSKPTESTTESSEAPTTPAKTSETKPSNVSSTSRKSTENVETSTSQ 2702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TPKKPAPKELAPTTTKGPT-STTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EDSLTVSVRIHELTISSENVPKESESTTTSSESSKPSQEPAGILTSTVVVPTSSVSLITA 2582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPA-----PT-----
DAKKRKPPANGLPKKKSAKESSSSEDSSSEEDEPPAKKRAQ----PAGGKKPVVKAVQPK 114
                                            DNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPR 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPK---TTPTPRKMTST 859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VVTASSIPSEEPILSSVTSSSTPRVRLITGTPDDLIVSVTVPSHGNRRQNITASSV---P
                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----TP--KGTAP--TTPK-----EPAPTTPKEPAPTTPKGTAPTTLKEPAPT 593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----QVTSTTTQDTTP-FKTTT-----KTTTTLAPKVTTTKKTITTTEIMNKP 802
                                                                                         12.5%; Score 626.5; DB ilarity 26.2%; Pred. No. 7e-20; Conservative 132; Mismatches 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -PKEPAPT-----TPKEPAP---
                                                                                                                                        DB 2;
                                                                                                393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              #text_change
                                                                                                                                                                                                                                                                                                                                                                                                                                    nucleolar phosphoprotein, xNopp
                                                                                                Indels
                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -TTPKETAPTTP 510
                                                                                                                                             990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24 - Sep - 1999
                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2865
                                                                                                  45;
```

```
A;Gene: GDB:TTN
A;Cross-references: GDB:127867; OMIM:188840
                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
Query Match 12.6
Best Local Similarity 23.3
Matches 288; Conservative
                                                                      position:
                                                                                                                                                                                                                                                                                                                                                                    704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYSPPVQP----PPVQKPPTPTYSPPVKPPPTQKPPT----PTYSPPIKPPPVKPPTPTY
                                                                                                                                                                                                                                                                                                                                                                 PPVKPPPVQVPPTPTYSPPIKPPPVQVPPTPTTPSPP
                                                                                                                                                                                                                                                                                                                                                                                          APKMTKETATTTEKTTESKITATTTQVTSTTTQDTTP
                                                                                                                                                                                                                                                                                                                                                                                                                      ---TPTYSPPIKPPPVQKPPTPTYSPPVKPPPVQLPP-TPTYSPPVKPPPVQVPPTPTYS
                                                                                                                                                                                                                                                                                                                                                                                                                                               DESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VHKPPTPTYSPPIKP----PPVHKPPTPTYSPPIKPPPVHKPPTPTYSPPIKPPPVHKPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KPAPTTPKETAPTTPKEPAPTTPKKPAPT-TPETPPPTTSEVSTPTTTK--EPTTIHKSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPPIKP-PPVHKPPTPTYSPPIKPPPIHKPPTPTYSPPIKPPPVHKPPTPTYSPPIKPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TTPKEPAPTTPKEPAPT-TPKGTAPTTLKEPAPT--TPKKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KETAPTTPKGTAPTTLKEPAPT - - TPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KLPPVKPPTPIYSPPVKPPPPVHKPPTPIYSPPVKPPPVHKPPTPTYSPPIKPPPVKPPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ELAPTTPEEPTPTTPEEPAPT----TPKAAAP-NTPKEPAPTTPKEPAPTTPKEPAPTTP 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTPIXSPPVKPPPVHKPPTPIXSPPVKPPPIQKPPTPTYSPPIKPPPLQKPPTYTYSPPI 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAP--TAPKEPAPT----TPKETAPTTP----KKLTPTTPEKLAPTTPEKPAPTT---PE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VKSPPVQKPPTPTYSPPIKPPPVQKPPTPTYSPPIKPPPVKPPTPTYSPPVKP-PPVHKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APTTPKEPSPTTTKEP---APTTPKEPAPT--TPKKPAPTTPKEPAPTTTKK 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KPAPT-TPKEPAPTTPKEPAP----TTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PPIKPPVHKPPTPIYSPPIKPPPVHKPPTPIYSPPIKPPPVHKPPTPTYSPPVKPPPVHK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TPKKP---APTTPKEPAPTTP----KEPTP--TTPKEPAPTTKEPAPT-TPKEPAPTAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YSPPIYP----PPIQKPPTPSYSPPVKPPPV---QMPPTPTYSPPIKPPPVHKPPTPTYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTPKAETTTKGPALTTP---KEPTPT-TPKEPASTTPKEPTTTTKSAPTTPKEPAPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PPTPIYSPPIKPPPVHKPPTPIYSPPVKPPPVQTPTPIYSPPVKPPPVHKPPTPTYSPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KSAPTTPKEPAPTTTKEPAP--TTPKEPAPTTTKEPAPTTTKSAPTTP----KEPAPT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PPPIQKPTYSPIYPPPIQKPPTPTYSPPIQKPPTPTYSPPIYPPPIQKPPTPT
                                                                    2q31-2q31
           12.6%; Score 632; DB 2; 23.3%; Pred. No. 2.7e-19;
 102;
 Mismatches
                                                                                                                      PIDN:CAA62189.1; PID:g1017427
442;
                                                                                                                                                                                                                                                                                                                                                                                            771
                                                                                                                                                                                                          ultrastructure
                       Length 7962;
Indels 406;
                                                                                                                                                                                                                                                               21-Jul-2000
                                                                                                                                                                                                         and
Gaps
                                                                                                                                                                                                       elasticity
                                                                                                                                                                                                                                                                                                                                                                                                                        703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           647
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              187
58;
```

```
B
                                    Q
                                                                      В
                                                                                                    Qy
                                                                                                                                         밁
                                                                                                                                                                    δÃ
                                                                                                                                                                                                           Вb
                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                             20
                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                  δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                        6595 IHEEEEFITEEEVVPVIPVK-VPEVPRKPVPEE-KKPVPVPKKKEAPPAKVPEVPKKPEE
                                                                                                                                                                                                                                                                                                                                                                                                            6482 KAVP-EEKVPVP-IPKKLKPPPPK--VPEEPKKVFEEKIHISITKREKEQVTEP---AAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5837
                            709 KPEMTTTAKDKTTERDL--RTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTT 766
                                                                                                                                                                                                                                                                                                                                                                        453 TPEEPTPTTPEEPAPTTPKAAAP--NTPKEPAPTTPKEPA-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10
                                                                                                                                                       KPAPKEL----APTTTKGPTSTTSDKPAPTTPKETAPT-TPKE---PAPTTPKKPAPTTP
EEEYVPEEEEFVPEEEVLPEVKPKVPVPAP-----VPEIKKKVTEKKVVIPKKEEAPPAK 6876
                                                                PVALPQEEEVLFEEEIVPEEEEVLPEEEEVLPE - EEEVLPEEEEVLPEEEEIPPEEEEVPP
                                                                                                                                                                                                      KVPVLIPKKEKP----PPAKVPEVPKKPVP-EEKVPVPVPKKVEAP-----PAKVPEVPK
                                                                                                                                                                                                                                                                                                       -----PTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTS
                                                                                                                                                                                                                                                                                                                                         VPMKPKRVVAEEKVPVPRKEVAPPVRVPEVPKELEPEEVAFEEEVVTHVEEYLVEEEEEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KRVVKEEKVSIEAPKREPOPIKEVTIMEEKERAYTLEEEAVSVOREEEYEEYEEYDYKEF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VSKKTVVEEKRFVAEEKLSFAVPQRVEVTRHEVSAEEEWSYSEEEEGVSISVYREEEREE 6181
                                                                                                                                   KPVPEKKVPVPAPKKVEAPPAKVPEVPKKLIPEEKKPTPVPKKVEAPPPKVPKKREPVPV
                                                                                                                                                                                                                                    TT-----SDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA--PTTPK 596
                                                                                                                                                                                                                                                                                                                                                                                                                                             KPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEK-----LAPTTPEKPAPTTPEELAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAKVPEAPKKIVPEKKVPAPVPKKEKVPPPKVPEEPKKPVPEKKVPPKVIKMEEPLPAKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----PTAPKK------PAPTTPKEP--PAPTTTKEPSPTTPKE----PAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PEIPKKKVPEERK-PVPRKEEEVPPPPKVPA--LPKKPVPEEKVAVPVPVAKKAPPPRAE 6121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VPAVHTKKMVISEEKMFFASHTEEEVSVTVPEVQKEIVTEEKIHVAVSKRVE-----P
                                                                                                 ETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EEYEPTEEYDQYEEYEEREYERYEEHEEYITEPEKPIPVKPVPEEPVPTKPKAPPAKVLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----PK-EPAP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TEKHMQITQEEKVLVAVTKKEAPPKARVPEEPKRAVPEEKVLKLKPKREEEPPAKVTEFR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTKSAPTTTKEP--APTTTKSAP---TTPKEPSPTTTKE------PAPTT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EEEAEVTEYEVMEEPEEYVVEEKLHIISKRVEAEPAEVTERQEKKIVLKPKIPAKIEEPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VPEVPKKPVPEEKIPVPVAKKKEAPPAKVPEVQKGVVTEEKITIVTQREESP----PPAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTP-KEPAPTTPKEPTPTT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PPKVPELPEKPAPEEVAPVPIPKKVEPPAPKVPEVPKKP-VPEEKKPVPVPKKEPAAPPK 6008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SAKETQSIEKTSAKDLAPT--SKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPAS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KKAPPPPKVIKKPVIEKIEKTSRRMEEEKVQVTKVPEVSKKIVPQKPSRTPVQEEVIEVK 5896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --TTPKEPTPTTIKSAPTTPKEPAP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KKKPTP----KPPV---VDEAGSGLDNGDFKVT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -PKITTAKPINPRPSLPPNSDTSKETSLT---VNKETTVETKETTTNKQTSTDGKEKTT 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ------PKEPA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -TTPKKPAPT - - TPKEPAPTTPKEPAPTTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TTTKSAPTTPKEPAPTTTKEPA-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --TPDTSTTQHNKVSTS
                                                                                                                                                                    648
                                                                                                                                                                                                      6702
                                                                                                                                                                                                                                                                          6652
                                                                                                                                                                                                                                                                                                                                                                                                              6534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                   708
                                                                                                                                                                                                                                                                                                                                             6594
                                                                                                                                                                                                                                                                                                                                                                                                                                                452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6064
                                                                                                                                                                                                                                                                                                           544
                                                                                                                                                                                                                                                                                                                                                                             490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              50
```

```
RESULT 11

T11622
extensin class 1 precursor - cowpea
C;Species: Vigna unguiculata (cowpea)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-1
C;Accession: T11622; S54155
R;Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
Mol. Plant Microbe Interact. 10, 95-101, 1997
A;Title: Rhizobia modulate root-hair-specific expression of extensin A;Reference number: Z17301; MUID:97155574
A;Accession: T11622
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Boolecule type: DNA
A;Boolecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
A;Description: A class of root-hair specific extensins involved A;Reference number: $54155 A;Accession: $54155 A;Status: preliminary A;Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δδ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                     A;Cross-references: EMBL:X91836; NID:g1015936; PIDN:CAA62943.
A;Experimental source: sub_species Red caloona
R;Arsenijevic-Maksimovic, I.; Broughton, W.J.; Krause, A.
submitted to the EMBL Data Library, April 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 뮍
                                                                                                                                                               A; Residues: 1-489 < ARS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                   1023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STTTALPFTTEQTVTTEEPTTAEKSTATQ----KPTTTQESVST---EKTSTTKKA---S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTPKKLTPTTPEKLA---PTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SHVVPKPTKKGTVKVTPKLELSFDEPTEIT-KAPHPGKLLEKKTYHFVLSDNFARYSEAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   E---PAPT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STPRQLPERWKAIVNKFKHNLEVLKEKKRLLKEKESTSTTGSDSSETTTVVAENIDEVTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTTKEPTSTTSDKPAP----TTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TTEE--PTTTDEPTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENDDYNHLDYNHYREAKEPTTTEESSTTEEVTTTEEPANTGNPPTTENPTTTEQPTSTAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                   HTTTDAAFVTATEASLNDGSDKKIIDEAQPTDEIRRA 1059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SKITATTTQ-VTST---TTQDTTPFKITTLKTTTLAPKVT---TTKKTITTTEIMNKPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTTA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TEKEKVVQTTPITTEKSTTQEETTTTTTTTEKTTSKTTTEKPTTSESATTETTTSEPST-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPE--TPPPTTSE-VSTPTTTKEPTTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TTTEEPTTTAIFAEASTGIITTDEETTSTTSTTPEITSTKE--IVTESAITQTSVSVVES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KTTPTPRKMTSTMPELN - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----INNTQISQPKPTDISKTDALSSLISGLIGSFTKAPMAPTI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TAKPKDRATNSKATTPKPQ-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VAASTTTTEPITTTEKSTTLETTPIEATTLNEVTGPAFVTGAPVDETTINTLELLSK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTTSESAAFITGESPENTALQSSSQKSEENESSAEKPGARRDFVPKKHKTTVKPAETTSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---TEST---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----PAPTTPKEPAP----TTTKKPAPTA--PKEPAPTT---PKETAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----TVDTSSATTEESSTAAETTTTSAE---TSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TPK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---ESSTTGKATTPELSTTSEETTTTELKITTE--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPAPTTPKKPAP - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --KKPAPKELAPTTTKGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTSRIAEA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -KPTKAPKKPTSTKKPKTMPRVRKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    - APKMTKETATTTEKTTE
                                                                                                                                                                                                                                                          extensin
                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----TS
                                                                                                                                                                                                                                                                                                                              17-Nov-2000
                                                                                                                                                PID:g1015937
                                                                           in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -TTPK
                                                                             rhizobium/legume
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GS
                                                                                                                                                                                                                                                        genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           670
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              982
```

inte

Conservative

```
A;Residues: 326-489 <AR2>
A;Cross-references: EMBL:X86030; NID:g791149; PID:g79
A;Cross-references: EMBL:X86030; NID:g791149; PID:g79
C;Genetics:
A;Gene: Ext26G
C;Superfamily: hydroxyproline-rich glycoprotein
C;Keywords: glycoprotein; hydroxyproline
F;1-23/Domain: signal sequence #status predicted
F;24-489/Product: extensin class 1 #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
C84672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ΩV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                A,Title: Sequence and analysis of chromosome A;Reference number: A84420; MUID:20083487 A;Accession: C84672
                                                                                                                                                                                                                                                                                       C;Accession: C84672
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. R;Lin, X.; Kaul, S.; Cronin, L.A.; Shea, M.; Vanaken, S.E.; Umayam, L.; Tallon, M.; Koo, H.; Moiffat, K.S.; Cronin, L.A.; Shea, M.; Vanaken, S.E.; Umayam, L.; Tallon, euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
                                                                                                                                                                                                                                                                                                                                                                                    hypothetical protein At2g27380 [imported] - Arabidopsis thaliana C;Species: Arabidopsis thaliana (mouse-ear cress) C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    B
                                                                                                                                   A;Cross-references:
C;Genetics:
                                                                                                                                                                        A; Molecule type: DNA
A; Residues: 1-761 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                               A;Map
                                                                                                                 A; Gene:
                                                                                                                                                                                                                A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Best Local Similarity Matches 217; Conserv
                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       179 PKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAP
                                                                                               position:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
nes 149; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       39
                                                                                                 At2g27380 position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PSPSPPPPYYYKSPPPPSPSPPPPYYYKSPPPPPS--PSPPPPYYYKSPPPPSPSPPPPYY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PKQTPPYYYNAPPYYYKSPPP-----PSPSPP-----PPPPYVHKYPPYYYKSPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----KSPPPPSPSPPPPYYYKSPPPPSPSPPPPYYYKSPPPPSPSPPPPYYYKSPPPPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        YYYKSPPPPSPSPYYYYKSPPPPSPSPPPPYVYKSPPPPSPSPPPPYYYKSPPPPSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 YKSPPPPSPSPPPPPYVYKSPPPPSPSPPPPYYYKSPP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PSPSPPPPYYYKSPPPPSPSPPPPYYYYSPPPPSPSPSPPPPPSPSPSPPPPPYYY-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PPPPYYYKSPPPS9SPPPPYYYKSPPPS9SSPPPPYYYKSPPPSPSPPPPYYYKSPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTP---KEPAPTAPKKPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELA 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GB:AE002093;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12.6%;
12.6%; Score 632; DB 2; 31.1%; Pred. No. 3.2e-20; bive 44; Mismatches 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 633; DB 2.
Pred. No. 2e-20;
33; Mismatches 29
                                                                                                                                                       NID: g5306260;
                                                                                                                                                                                                                                                                             N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PID:g791150
                                                                                                                                                                                                                                                                             of
                                                                                                                                                             PIDN: AAD41992.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <SIG>
                                                                                                                                                                                                                                                                             the plant Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   <MAT>
                                        Length 761;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -PPSPSPPPP
                                                                                                                                                             GSPDB:GN00139
        82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    475
        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 307
        33;
```

```
8
                                                                                        S
                                                                                                                                    망
                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Map position: 1921-1923
A;Introns: 20/1; 62/3; 1165/3; 1184/2; 1230/1; 1270/3; 1320/3
C;Superfamily: polymorphic epithelial mucin
C;Keywords: alternative splicing; duplication; glycoprotein; phosphoprotein; polymorphic epithelial mucin
C;Keywords: alternative splicing; duplication; glycoprotein; phosphoprotein; polymorphic epithelial
C;Keywords: alternative splicing; duplication; glycoprotein; phosphoprotein; polymorphic epithelial
C;Keywords: alternative splicing; duplication; glycoprotein; phosphoprotein; polymorphic epithelial
C;Keywords: alternative splicing form A *status predicted <PREA>
F;1-1344/Product: mucin 1 precursor, splice form B *status predicted <SIGB>
F;1-19,29-32/Domain: signal sequence *link PREB *status predicted <SIGB>
F;1-19,29-134/Product: mucin 1 precursor, splice form B *status predicted <PREB>
F;1-19,29-212,103-1344/Product: mucin 1 precursor, epithelial tumor antigen splice
F;1143-1314/Region: 20-residue repeats (GSTAPPAHGVTSAPDTRPAP)
F;1143-1314/Region: mucin 1 carboxyl-terminal non-repetitive
F;1245-1272/Domain: transmembrane *status predicted <TRM:
F;1046,1064,1118,1144,1222/Binding site: carbohydrate (Asn) (covalent) *status predicted
F;1213/Binding site: phosphate (Tyr) (covalent) *status predicted
                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C;Genetics:
A;Gene: GDB:MUC1; PUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Cross-references:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                             572
                                                                                                                                       605
                                                                                                                                                                                 527
                                                                                                                                                                                                                           545
                                                                                                                                                                                                                                                                       487
                                                                                                                                                                                                                                                                                                                 485
                                                                                                                                                                                                                                                                                                                                                           448
                                                                                                                                                                                                                                                                                                                                                                                                       425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            343 TKSAPTTPKEPSPTTTKEP-----APTTPKEPAPTTPK----KPAPTTPKEPAPTTPK-
                                                                                                                                                                                                                                                                                                                                                                                                                                                 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           134 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTP------KE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             77
  TTKGPTSTTSDKPAP -- TTPK -- -- ETAPTTPKEPAPTTP-
                                     AHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP-
                                                                                                                           HGVTSAPDTRPAPGSTAP-PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPP
                                                                                                                                                                                                                                                                                                     HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                               HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                               ---ELAPTTPEEPTPTTP-----EEPAP--TTPKA----AAPNTPKEPAPTTP--
                                                                                                                                                                                                                                                                                                                                                                                         HGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                T-SAPDT--RPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KPAPTTPKEPAPTTPK----EPAPTTTKEPSPTTPKEPAPTT---TKSAPTTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --APTTPKEPTPTTP------KEPAPTTKEP-----APTTPKEPAPTAPK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSAPDT---RPAPGSTAPPAHGVTSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LTVLTATTAPKPATVVTGSGHASSTPGGEKETSATORSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LTVNKETTVETKETTTINK--QTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKYLAKPT
                                                                                                                                                                                                                                                                                                                                                                                                                                 ---EPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDT--RPAPGSTAPPAHGVT-SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP-----KKPAPTTPKEP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTPTT-----IKSAPTTPKEPAPTTT-----KSAPTTPKEPAPTTTKEP-----A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVTSAPDNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       302;
                                                                                                                                                                     -----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKG----TAPTTPKEPAPTTP-
                                                                                                                                                                                                                                                           -----KEPAPTTPKEP-----APTT---PKETAPTTPKGT-APTTLKEPAPTTP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDB:120705; OMIM:158340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 666.5; DB 1; Length 1344;
Pred. No. 1.9e-21;
70; Mismatches 432; Indels 261;
·-KKPAPTTPET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                             663
                                                                                                                                                                       571
                                                                                                                                                                                                                                                                                                                                                                                              484
                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                                                                                                                    486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                                                                                                                      447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polymorphi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     form
```

```
Š
                                                                                    ġ
                                                                                                                                                             Дb
                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Nolecule type: DNA
A; Molecule type: DNA
A; Residues: 1-1229 <FUL>
A; Cross-references: EMBL:U67956; PIDN:AAB07691.1; GSPDB:GN00028; CESP:F16F9.2
A; Cross-references: Strain Bristol NZ; clone F16F9
          밁
                                                                                                                      Qy
                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                          Ωy
                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                               DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Description: The sequence
A;Reference number: Z20071
A;Accession: T25697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  submitted to the EMBL Data Library, August 1996 A; Description: The sequence of C. elegans cosmic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hypothetical protein F16F9.2 - Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Species: Caenorhabditis elegans C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000 C;Accession: T25697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Introns: 35/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R; Fulton,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1012 DTRPAPGSTAPPAHGVTSAPDNRPALGSTAPPVHNVTSASGSASG 1056
        382
                                                                                                                                                             273
                                                                                                                                                                                                  201
                                                                                                                                                                                                                                         235
                                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                                                                                                                                               119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 704
                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    723
                                                                                                                                                                                                                                                                                                                                                                                                                         33 KVTT-PDTSTTQ---HNKVSTSPKITTA-KPINPRPSLPPNSDT-----SKETSLTVN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity les 307; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10
SIEIPITTVPLVETTSTTSTASKESDGFHTTLKLKVTTADSDSTESATTVKPFNEETTTK
                                                                                                                                                                                                                                                          KGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPT 200
                                                                                                                                                                                                                                                                                                                                   KETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTT 140
                                                                                                                                                                                                                                                                                                                                                                                     ----PSPTTPKEPAPTTTKSAP-----TTTKEPAPT----TTKSAPTTPKEPSPTTTK 359
                                                                                                            PKEP-TPTTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKE- 314
                                                                                                                                                     TTEEPTPTTTEE---VTTTEAETSTTTSSETSTEK---PTTPLIDNKIAGPATGK---PET
                                                                                                                                                                                TTKEPAPTTPKEPAPTTTKSAPTTPKEPAPTTP----KKPAPTTPKEPAPTT 256
                                                                                                                                                                                                                             TAIAEST-----TTVIAEVSTTTEEPTTT------AESTTKKSTT----KAPA 272
                                                                                                                                                                                                                                                                                                           PETLVSTDSSTST-EQTSPDNTTEIASPMETNTTTEATTTSVEPSVSTLAS---EDETTV
                                                                         THEPVTGTTPN--FDTATETPEVAKSEDKMTLSKTAATETTQQTTEVT-DGPEKETTKNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAP----DTRPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TPAA-----TKPEMTTTAKOK---TTERDLRTTPETT----TAAPKMTKETATT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              {\tt APPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGST}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ELNPT-SRIAEAMLQTTTRPNQTP--NSKLVEVNPKSEDAGGAEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGSTAPPAHGVTSAPDTRPAP--GSTAPPAHGVTSAPDTRP---APGSTAPPAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TEKTTESKITATTTQVTSTTTQDTTP--FKITTLKTTTLAPKVTTTKKTITTT--EIMNK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGS--T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PPPTTSEVSTPTTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                13.2%; Score 659.5; DB 2; 29.0%; Pred. No. 3.5e-21; ative 78; Mismatches 339;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cosmid F16F9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels 333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                     324
                                                                                                                                                                                                                                                                                                             234
                                                                                                                                                                                                                                                                                                                                                                                       178
                                                                                                                                                                                                                                                                                                                                                                                                                           80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                52;
```

```
δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Вþ
                               N:Contains: mucin 1 precursor, epithelial tumor antigen splice f C;Species: Homo sapiens (man) C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change C;Accession: A35175; B35175; A35886; A35887; S10572; S40293; A36 R;Ligtenberg, M.J.L.; Vos, H.L.; Gennissen, A.M.C.; Hilkens, J. J. Biol. Chem. 265, 5573-5578, 1990
                                                                                            mucin 1 precursor, repetitive splice form A [validated] - human N; Alternate names: breast carcinoma-associated DF3 antigen; correctic mucin; polymorphic epithelial mucin (PBM) N; Contains: mucin 1 precursor, epithelial tumor antigen splice 1 C; Species: Homo sapiens (man)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δõ
                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                              γO
                                                                                                                                                                                                                                                                                                                 밁
> >
                                                                                                                                                                                                                                                  뫄
               . Biol. Chem. 265,
Title: Episialin,
                                                                                                                                                                                                                                                                                                                                                                                     1030
                                                                                                                                                                                                                                                      1105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      449
                                                                                                                                                                                                                                                                                                                                                                                                                                                      976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P----APTTPKEPAPTTPKKPAPTTPKEPAP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EPAPTTPKEPTPTT----PKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MGAATPPQTPSVGAATTP-QPSPM----GAATTLMSPMGAATTPQ---PSPMGAVTTQPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     \tt GGATQMSPMGAANTQMSPMGATTTQMSPMGAAATTQPSPMGAAATQVTATSAGNTMQVSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      {\tt TLMAPMGAANTQMAPIGAAGAAQTAPMGAAHTHVSPMGAGGATQMSPTGAANTHMSPIGA}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----PAPTTPKEPAPTTPKEPAPTTPKETAPT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LPSSPVASAMHAKVTPRPLPASPVPMAASPASLGPDAARVALATNAASPGAKPEAAGGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KPVPLMALTPQPVTAQMVTQLAATKPSPIVPKASPK-ALMTPPPPPPGLPRALAAAKLLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PGVPNAATDPQKPPPTPQSVPSAVTEPKPQPRAAPPPSNEATPAVPSPSPNLKSPLPTIP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PVTKAATTTNATPPPQPIPKAATTTTATPVTPQQPIPKAGTDAAPPPAVPKAPSDGRAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P---APTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAP---TTTKSAPTTPKEPSPTTTKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -TAGAVPKASTGTTPAAAPQQPVP---KAAPVTPPSPQQAVPRAATAAA----APVTPQQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KG--TAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEP--APTTPKKPAPKELAPTTTKGP
                                                                                                                                                                                                                                                                                     EVNPKSEDAGGAEGE
                                                                                                                                                                                                                                                                                                                                                                                                                     TPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMTTTAKOKTTEROLRTTPETTTAAPKMT-KETATTTEKTTESKITATTTQVTSTTTQDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PPPPTIPPSPTAQTSPQPMSKSPPPDPPKAPSAAAQTSPAAHVANASPGV-TAVSPA---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PTTSEVSTPTTTKEPTTIHKSPDESTPEL--SAEPTPKALENSPKEPGVPTTKTPAATKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PMAATNTTQPPPMAASTPQSTPMGAATTTQSPPMGATTTQSPPMGASTPQAPPTVAGSPT
                                                                                                                                                                                                                                                                                                                                                    KKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                      IGVTEASPSADGARLSPGPTAATDGPKASPAATADVTEAATD--VTAAATAVPA----EA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -TSTTSDKP-APTTPKET---APTTPKEP---APTTPKKP-APTTPETPP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ----EPAPTTPK---ETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELA-----
 a carcinoma-associated A35175; MUID:90202794
                                                                                                                                                                                                                                                                                                                                                                                     ---EEPTPTTP----EEPAPTTP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TLKEPAPTTPKKPAPKELAPTTTKEP----
                                                                                                                                                                                                                                                                                       904
                    mucin,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -TTPKEPAPT-TTKKPAPTAP
                    s.
                   generated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -TPKGTAPT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -TSTTSDKPAPTTP
                                                                                                                                                      core
                                                      А36735;
J.
                    bу
                                                                                                                     form; mucin 1
                                                                                       02-Jun-2000
                                                                                                                                                      protein
                    ω
                    polymorphic
                                                                     PX0066;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            450
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               396
                                                                                                                                                                                                                                                                                                                                                                                           1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          975
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           480
                                                                                                                                                       KP39;
                                                                                                                     precursor,
                                                                                                                                                      episiali
                      gene
                      enc
                                                                       SS
```

```
A; Residues: 1-19,29-155, 'p',157-175, 'p',177-182, 'A',184-212,1033-1037, 'A',1039-1344
A; Residues: EMBL:X52229; NID:g37053; PIDN:CAA36478.1; PID:g37054
R; Abe, M.; Siddiqui, J.; Kufe, D.
Biochem. Biophys. Res. Commun. 165, 644-649, 1989
A; Title: Sequence analysis of the 5' region of the human DF3 breast carcinoma.
A; Reference number: A36735; MUID:90088473
A; Accession: A36735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Molecule type: mRNA
A:Residues: 1-19,29-992,1033-1344 <GEN>
A:Residues: 1-19,29-992,1033-1344 <GEN>
A:Residues: In-19,29-992,1033-1344 <GEN>
A:Ross-references: GB:JO5581; NID:g188869; PIDN:AAA59876.1; PID:g188870
A:Construction of the tandemly repeated
A:Note: GenBank entry HUMMUCAB includes one copy of the tandemly repeated
A:Lan, M.S.; Batra, S.K.; Qi, W.N.; Metzgar, R.S.; Hollingsworth, M.A.

Biol. Chem. 265, 15294-15299, 1990
A:Title: Cloning and sequencing of a human pancreatic tumor mucin cDNA.
A:Reference number: A35887; MUID:90368716
A:Accession: A35887; MUID:90368716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: GB:M32738; GB:J05288; NID:g182121; PIDN:AAA35804.1; A;Experimental source: splice form A A;Note: GenBank entries HUMEPISIAl and HUMEPISIA2 present only the amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Accession: A35175
A; Molecule type: mRNA
A; Residues: 1-952,1033-1344 <LIG1>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               R; Cross reconstruction of the EMBL Data submitted to the EMBL Data submitted to number: $40293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Residues: 1-19,29-952,1033-1344 <LIG2>
A; Cross-references: GB:M32739; GB:J05288;
A; Experimental source: splice form B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Note: GenBank entries
A; Accession: B35175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Molecule type: mRNA
A;Residues: 1-19,29-1109,'S',1111-1339,'A',1341-1344 <LAN>
A;Cross-references: GB:J05582; NID:g189598; PIDN:AAA60019.1;
A;Note: GenBank entry HUMPANMU contains four fewer copies of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Title: Molecular cloning and expression A;Reference number: A35886; MUID:90368715 A;Accession: A35886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;Gendler,
J. Biol. C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                 A;Molecule type: mRNA
A;Residues: 1-142,'Q',144-162,'Q',164-168 <ABE>
A;Cross-references: EMBL:M31823; NID:g181542; PIDN:AAA35757.1; PID:g181543
A;Cross-references: EMBL:M31823; NID:g181542; PIDN:AAA35757.1; PID:g181543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: mRNA
A; Residues: 1-19, 29-155, 'P', 157-175, 'P', 177-182, 'A', 184-212, 1033-1344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Title: Human epithelial tumor antigen cDNA A; Reference number: S10571; MUID: 90276413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R;Wreschner, D.H.; Hareuveni, M.; Tsarfaty, I.; Smorodinsky, Eur. J. Biochem. 189, 463-473, 1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Note:
A;Title: Tyrosine phosphorylation of the MUC1 A;Reference number: S51026; MUID:95080414 A;Contents: annotation
                                                                                                                                                                                                                                                           A; Title: A novel core protein as well as polymorphic A; Reference number: JX0235; MUID:93123189
                                                                                                                                                                                                                                                                                                                          R;Masuzawa, Y.; N
J. Biochem. 112,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Accession: S10572
                                                                                   A; Experimental source: gastric carcinoma cell R; Zrihan-Licht, S; Baruch, A; Elroy-Stein, FEBS Lett. 356, 130-136, 1994
                                                                                                                                                                             A; Molecule type: mRNA
A; Residues: 998-1011, 'ES', 1014-1017; 1018-1032, 'T', 1034-1037; 1038-1057
                                                                                                                                                                                                                                   A; Accession: PX0066
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cross-references: EMBL: X52229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Status: not compared with conceptual translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Biol. Chem. 265, 15286-15293, 1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GenBank entries HUMEPISIB1 and HUMEPISIB2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S.J.; Lancaster,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1-19,29-952,1033-1344 <LIG2>
                                                                                                                                                                                                                                                                                                                       .; Miyauchi, T.; Hamanoue, M.; Ando, 12, 609-615, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Α.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NID: 937053
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Taylor-Papadimitriou,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NID: g182126;
                                                                                                                          0.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequences. Differential splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     human tumor-associated
                                                                       breast cancer membrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              present of triou, J.;
                                                                                                                                                                                                                                                                                                     epithelial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIDN:AAA35806.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              only the amino-and .; Duhig, T.; Peat,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Duhig,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PID:g189599
the tandemly
                                                                                                                                                                                                                                                                                                        mucin carry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        repeated, M.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Horev, J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino-and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polymorphic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     carcinoma-associa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PID:g182124;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PID:g182129;
                                                                                                                                   D.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  repeated
                                                                                                                                                                                                                                                                                                                                                                s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence
                                                                                                                                                                                                                                                                                                           peanut
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zaretsky
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   carboxyl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            carboxyl
                                                                          cytokin
```

A;Note: undetermined tyrosine residues in the carboxyl-terminal non-repetitive C;Comment: This protein is length polymorphic. Individuals may have between 21 partial repeats. The repeat shown is defined by SmaI nuclease sites. C;Comment: Serine and threonine residues in the tandem repeat domain are extens C;Comment: For an alternative splice form without a tandem repeat domain, see F

are extensively main, see PIR:S4

regi and ion 12 proteins

```
A; Accession: A26877
A; Molecule type: DNA
A; Residues: 1-242 <YAM>
A; Residues: 1-242 <YAM>
A; Residues: 1-242 <YAM>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Cross-references: EMBL:M16165; NID:g172523; PIDN:AAA35015.1; PID
A; Molecule type: DNA
A; Residues: 762-1331 <YAM>
A; Cross-references: EMBL:M16165; NID:g172523; PIDN:AAA35015.1; PID
R; Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.
A; Title: Similar short elements in the 5' regions of the STA2 and (A; Reference number: S27281; MUID:89031230
A; Accession: S27281
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; First Company (A)
A; Residues: 1-31 < PAR>
A; Cross-references: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16166; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16166; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16164; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Residues: 1-31 < PAR>
A; Cross-references: EMBL:M16166; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Cross-references: EMBL:M16166; NID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Cross-references: MID:g172522; PIDN:AAA35014.1; PID
A; Molecule type: DNA
A; Molecule type
                                              Ş
                                                                                              밁
                                                                                                                                       S
                                                                                                                                                                                        В
                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A:Map position: 9R
C:Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; gluca C:Keywords: glycosidase; hydrolase; polysaccharide degradation; F:5-21/Domain: transmembrane #status predicted <TMI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÔ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Gene: SGD:MUC1; STA2; MAL5; DEX2; SGD:S0001458
A;Cross-references: MIPS:YIR019c; SGD:S0001458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Status: nucleic acid sequence
A:Molecule type: DNA
A:Residues: 1-1367 <- CLAMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A: Cross references: EMBL:X13857; NID:g4551; PIDN:CAA32069.1; R: Lambrechts, M.G.; Bauer, F.F.; Marmur, J.; Pretorius, I.S. Proc. Natl. Acad. Sci. U.S.A. 93, 8419-6424, 1996
A:Title: Mucl. a mucin-like protein that is regulated by Mssi A; Reference number: JC6123; MUID:96323237
A:Accession: JC6123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Cross-references: GB:U30626; NID:g1304386; PIDN:AAC49609.1; PID:g1304387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;1350-1366/Domain: transmembrane #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                447
                                                                                                                                            391
                                                                                                                                                                                             584
                                                                                                                                                                                                                                                                                            525
                                                                                                                                                                                                                                         337
                                                                                                                                                                                                                                                                                                                                          285
                                                                                                                                                                                                                                                                                                                                                                                           465
                                                                                                                                                                                                                                                                                                                                                                                                                                         239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hes 280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  35
       EELAPTTPEEPTP---TTPEEPAPT-TPKAAAPNTPKEPAP----TTPKEPAPTT----
                                                                                                                           KEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAP----TTPEKPAPTTP
                                                                                                                                                                                                                                                                                                                        PTAPKKPAPTTPKEPAPTTPKE-----PAPTTTKEPSPTTPKEPAPTTTKSAPTTTK
                                                                            --PSSSTTESSSAPVPTPSSSTTESSSAP-VPTPSSSTTESSSAPVTSSTTESSSAPVTS
                                                                                                                                                                                                                  EPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTT-----PKEPAPTTP
                                                                                                                                                                                                                                                                          APAPTPSSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSSTPVTSSTTESSSAPVPT-
                                                                                                                                                                                                                                                                                                                                                                                                                       TTPKKPAPTTPKEPAPTTPKEPTP---TTPKEPAPTTKE-----PAPT----TPKEPA
                                                                                                                                                                           -PSSSTTESSSAPVPTPSSSTTESSSAPAPTPSSSTTESSSAPVTSSTTESSSAPVPT--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSSAPVTSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEPAP----TTTKEPAP----TTPKEPAP----TTTKEPAP----TTTKSAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSSTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LTTPKE-----PTPTTPKEPASTTPKEPTPTTIKSAP----TTPKEPAPTTTKSAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TKKKTTTSK-TCT---KKTTTPVPTPS-SSTTESSSAPV-----PTPSSSTTESSSAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETT--TKGPA 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STSESSTTTSSTSESSTSSSTTAPATPTTTSCTKEKPTPPTTTSCTKEKPTPPHHDTTPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTPDTSTTQHNKVSTSPKITTAKPINP------RPSLPPNSDTSKETSLTVNKETTVE
                                                                                                                                                                                                                                                                                                                                                                         TTESSSAPVTSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 762.5; DB 1;
Pred. No. 1.7e-25;
L5; Mismatches 441;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    regulated by Mss10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <TM2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 1367;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STA2 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               glucan 1,4-alpha-glucosidase
tion; transmembrane protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PID: 94552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PID:g172526
A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PID:g172525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SGA genes from Saccha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 critical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                           336
                                                                                                                                                                              640
                                                                                                                                                                                                                         390
                                                                                                                                                                                                                                                                             583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       404
                                                                                                                                                                                                                                                                                                                                                                           524
                                                                                                                                                                                                                                                                                                                                                                                                                          284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pseudohy
Š
                                                В
                                                                                            Ş
                                                                                                                                                  밁
                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                              ₽
                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
```

```
high molecular mass nuclear antigen - chicken (fragment) c;Species: Gallus gallus (chicken) C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text C;Accession: T18535 R;Shimada, K.; Harata, M.; Mizuno, S. J. Cell Sci. 110, 3031-3041, 1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Cross-references:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: mRNA
A; Residues: 1-1151 <SHI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: T18535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Title: A nuclear matrix-associated high A; Reference number: Z18955; MUID: 9803440
                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches . 284; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1047
        204
                                              278
                                                                                                                                                                    123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1167 VPTTYV----
                                                                                                                                                                                                            171
                                                                                                                                                                                                                                                                                           113
                                                                                                                                                                                                                                                   65
                                                                                                                                                                                                                                                                                                                   14 TPKP-----PVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTS--PKITTAKPINPRPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   864 NPTSRIAEAMLQTTTRPNQ--TPNSKLVEVNPKSED---AGGAEGET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       987 VCSTGTNSAGETTSGCSPKTITTTVPCSTSPSETASESTTTSPTTPVTTVVSTTVVTTEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         758
                                                                                                                                                       APTSKVLAKPTPKAETTTKGPALTTPKEPTP-TTPKEPASTTPKEP----TPTTIKSAPT
                                                                                                                                                                                                                                                                              TPPPSQGPAGTPPPSQGAAGAPKGDGTAQPSGTKSGADGKPAAQDVPKATTA-ATEARP-
EPAPTTPKEPAPTTKEPAPTTTKSAPTT-
                                AEAKPVPATASLMATKVTAEAKPAPSPSVPKATTDTKAVTATAPKAGPDVKPAVAVCAEA
                                                                                                                 KPAAAAVAKEAKAVTAAAAAPKATAEAKPAPVTSPTIPCSSAEAKPLTAASPTASKA--T
                                                                                                                                                                                                                                    LPPNSDTSKETSLTVNKETTVETKETTTTNK--QTSTDGKEKTTSAKETQSIEKTSAKDL 122
                                                                                                                                                                                             ----ASAASPTVPKATAEATAVTAASQSAPKAATDAAAVTAA----SQSAPKATV-EV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGCSPKTVTTTVPCSTGTGEYTTEATTLVTTAVTTTVVTTESSTGTNSAGKTTTGYTTKS 1166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPEL 863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STSTKPGGEITTTFVTKNIPTTYLTTIAPTPSVTTVTNFTPTTITTTVCSTGTNSAGETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TTPFK-----ITTLKTTTLA--PKVTT----TKKTITTTEI---MNKPEET- 805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---TAKDKTTERDLRTTPET-TTAAPKMTKETATTTEKTTESKITATTTQVTST--TTQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TTPSTTTITTYCSTGTNSAGETTSGCSPKTYTTT-----VPTTTTTSVTTSSTTTITTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ITSSAPSSIPESSTTESESTGTT----VTPSSSKYPGSQTETSVSSTTETTIVPTKTTTSV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTKGP----TSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPET--PPPTTSEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ITSSAPSSTPFSSSTESSSVPVPT----PSSSTTESSSAPVSSSTTESSVAPVPTPSSSSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSVAPVPTPSSSSN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTESSSAPVTSSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----PKEPAPTTPKETAPT-TPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTST
                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL: D88440; NID: d1177138;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTLAPSAPVTPATNAVPTTITTTECSAATNAAGET
                                                                                                                                                                                                                                                                                                                                                                                         13.4%; 25.9%;
                                                                            ---TKSAPTTPKEPAPTTTK----
                                                                                                                                                                                                                                                                                                                                                                  100; Mismatches 419;
                                                                                                                                                                                                                                                                                                                                                                                   Score 671; DB 2;
Pred. No. 1.1e-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         molecular mass
  PKEPA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                             PID:d1025045; PIDN:BAA24137.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  #text_change
                                                                                                                                                                                                                                                                                                                                                                                                    Length 1151;
                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       905
                                                                              . . . . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   antigen, HMNA,
                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                         177
                                                                                                                                                                                                219
                                                                                                                                                                                                                                                                                 170
                                                                          203
                                                                                                                   277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           757
                                                                                                                                                                                                                                                                                                                                                             55;
```

of.

Qy	Оу	Фр	Оу	Qy Db	Qu Be	A; Ge A; Ma A; In A; No C; Ke	A; Re A; Cr C; Ge	A; Ac	R; Yo Gene A; Ti A; Re	C;Sp	T3082 nasce N;Alt	RESULT.	ОУ	DЪ	Qy	DЬ	Qy	₽ 5	} {	pb Oy	Db	Qy	Db
9 222 APTTTKSAPT	Y 172 IKSAPTTPKE-PAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEP 221 	145 LTT	85 VETKETTTUKOTSTDGKEKTTSAKETOSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPA	34 806	qe	Gene: Naca Map position: 10 Map position: 10 Introns: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3 Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding activ Note: differential splicing; DNA binding; transcription factor	Residues: 1-21 Cross-referenc Genetics:	ACCESSION: "INDXO Status: preliminary; translated from GB/EMBL/DDBJ MC10010 trop. DNA	R;Yotov, W.V.; St-Arnaud, R. R;Yotov, W.V.; St-Arnaud, R. Genes Dev. 10, 1763-1772, 1996 A;Title: Differential splicing-in of a proline-rich exon converts alphaNAC into a muscle A;Reference number: 220889; MUID:96312450		0826 0826 scent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse Alternate names: alpha-NAC protein	SHT-III 6	682 SAEPTPKALENSPKEPGVPTTKTPAATKPENTTTAKDKTTERDLKTTEET /31	~	681	1001 MSSPPPDEVKSPPPPAPVKSPPPPAPVSSPPPPVKSPPPPAPVSSPPPV 1057	564 -KEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPT 622	PPPTPVSSPPPAPKSSPPPAP		459 PTTPEEPAPTTPKAAAPNTPKE-PAPTTPKEPAPTTPKEPAPTTPKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVKGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTVGTAPTTTTTTTTTT			769 PPAPLSSPPPAPQVKSSPPPVQVSSP-PPAPKSSPPLAPVSSPPQVEKTSPPPAPLSSPP 827
No. No.	U # # #	> > > > > > > > > > > > > > > > > > >	, 2000 2000	1 2 4 5 5	: E:	0 B 0	D	D£ .	S H S	Db	Qy Qd	, Q	Db 43		, Oy	DЪ	Qy	Db	Qy	Db		, Q	Db

```
밁
                                                                                            문
                                                                                                                                     Ϋ́
                                                                                                                                                                                      В
                                                                                                                                                                                                                                    ¥
                                                                                                                                                                                                                                                                                       용
                                                                                                                                                                                                                                                                                                                                                                                      Ř
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1323 KKTPKTAVPKETSAPSEGVTAVPLEIPPSPRKAPKTAAPKETPAPS--PEGATTAPVQIP 1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1270 LTPKGAPNALAE-SPASPKKVPKTAAPEETSTTP-----SPQKIPKVAGPKEASATPPS 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1150 QKAPKTAVPKQIPTPEDAVTILAGSPLSPKKASKTAAPKEAPATPSVGVIAVSGEISPSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1090 GGVTAVPPETSLPPKETPQNATPNESLAASSQKRSPKTSVPKETPPGGVTAMPLEIPSAP 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1033 APKTAVSKEIPSKGVTAVPLEISLPLKETSKSATPGEKSASSPKRSPKTAGPKE---TPP
1753 DSHISPVS-DACSTGTTTP 1770
                                                                                                                                                                                                                                                                                     1612 T--SLAQTAPPSLOKAPSTTIPKENLAAPAV----LPVSSKSPAAPARASASLSPATAAP 1665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  507
                                                                                                                                     828 APKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSK 887
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TKEPAPTTPKEPAPTTPKKPAPTT--PKE------PAPTTPKEPAP-----T 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TKEPSPTTPK-----EPAPTTTKSAPTTTKEPAPTTT-KSAPTTPK------EPSPTT 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTT--PKEPAP----TTPKEPAPTT 311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEK-PAPTTPEELAPTTPE 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKTSKTAAPKENSATLPPKRSPKTAAPKETPATSSEGVTAVPSEISPSPPTPASKGVPVT
                                                                                                                                                                                        QTAPKEATTIPSCKKAAATETPIETSTAPSLEGAPKETSE----TSVSKVLMSSP-----
                                                                                                                                                                                                                                    DTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTK 827
                                                                                                                                                                                                                                                                                                                                                                                    KNPSSHKKTSKTIELKEAPATLPPSPTKSPKIPSSKKAPRT--SAPKEFPASPSIK-PVT 1611
                                                                                                                                                                                                                                                                                                                                                                                                                                  TSEVSTPTTTK-----EPTTIHKSPDESTPELSAEPTPKALENSPKE-PGVPTTKTPAA 707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKELAPTITKGPTSTTSDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP-----PPT 654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VGPKETSLEGATAVPLEIPPSHKKAPKTVDPKQVPLTPSPK-DAPTTLAE-SPSSPKK-A 1496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PSPRKGSKKAGSKE-TPTTPSPEGVTAAPLEIPISSKKTSKMASPKETLVTPSSKKLSQT 1439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EPTPTT------PEEPAPTTPKAAAPNTPKEPAPTTPKE-PAPTTPKEPAPTTPKETA 506
                                           LVEVNPKSEDAGGAEGETP
                                                                                                                                                                                                                                                                                                                                   TKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQ 767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKTAAPPSER-VTTVPPEKPA-TPQKASGTTASKVPVPAETQEVAVSSRETPVTPAVPPV 1554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----TTSDKPAPTTPKGTAPTTPKEPAPTTPKG-PAPTTPKGTAPTTLKEPAPTTPKKPA 599
                                                                                               -PKKASSSKRASTLP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTTPKGTAPTTLKEPAPTTP----
                                                                                                 -ATTLPSLKEASVLS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ---KKPAPKELAPTTTKEPTS--
                                                                                                      -PTATSSGK 1752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              544
                                                                                                                                                                                                     1716
```

```
$48478
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - yeast (Saccharomyces cerevisiae)
N;Alternate names: extracellular glucoamylase; mucin-like protein MUC1; protein YIR01
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Nov-1999
C;Accession: $48478; A26877; B26877; S27281; JC6123
R;Rowley, K.
submitted to the EMBL Data Library, October 1994
A;Reference number: $48478
A;Accession: $48478
A;Accession: $48478
A;Accession: $48478
A;Accession: $48478
A;Accession: $48478
A;Molecule type: DNA
A;Residues: 1-1367 <ROW-A;Residues: 1-1367 <ROW-A;Residues: 1-1367 <ROW-A;Residues: 1-1369 <ACCESSION: $48478
A;Cross-references: GB:Z47047; EMBL:Z38061; NID:g603997; PID:g763364; GSPDB:GN00009;
A;Title: Gene fusion is a possible mechanism underlying the evolution of STA1.
A;Reference number: A91831; MUID:87194600
```

```
S
                                                                                 밁
              밁
                                                                                                           Ş
                                                                                                                                                D
                                                                                                                                                                               Ş
                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                                                                                                                                                                                                                                                              망
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Molecule type: DNA
A;Residues: 1-1274 <LEI>
A;Cross-references: EMBL:U46675; NID:gl166613; PID:gl166621; PIDN:AAB52641.1; GSPDB:GN0Q
A;Experimental source: strain Bristol N2; clone F35A5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R:Leimbach, D.
submitted to the EMBL Data Library, January 1996
A:Description: The sequence of C. elegans cosmid F35A5.
A:Reference number: Z18485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: CESP:F35A5.1
A;Map position: X
A;Introns: 1272/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hypothetical protein F35A5.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Status: preliminary; translated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genetics:
919 K-KEPEKPKD-APKVAAKPRDPSPKKAVPE--KEPAKVAAKPRDLSPKKAIPIPANTQEA
                                                                                                                497
                                                                                                                                                  908
                                                                                                                                                                                 443
                                                                                                                                                                                                                    751
                                                                                                                                                                                                                                                        389
                                                                                                                                                                                                                                                                                         692
                                                                                                                                                                                                                                                                                                                                                            633
                                                                                                                                                                                                                                                                                                                        334
                                                                                                                                                                                                                                                                                                                                                                                             284
                                                                                                                                                                                                                                                                                                                                                                                                                                  573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               394 PKKAAPAVKPRDSSPKKATPLQADPKAQEVPPTPVKNPVKKYKPPWEVDDEDPVEEVKQP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         / Match 16.0%; Score 802; DB 2; Length 1274;
Local Similarity 28.0%; Pred. No. 3.5e-27;
hes 290; Conservative 102; Mismatches 414; Indels 230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51
                                   KGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPT----TPKKPAP-----KEL
                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEP----APTTPKKPAPTTPKE----PAPTTPKEPTPTTPKEPAPTTKEPAPTTPK-EP
                                                                                                  PAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT-TTKEPTSTTSDKPAPTTP
                                                                                                                                     PAEPVNVPEPEKKTPVLAKKTPVKPRDPSPKKAVPAKPSTKTDAPPVSVKKPEPVSKPKE
                                                                                                                                                                                                         T--EVPPAVVKKPEPVAKSRDPSPKKAK-AEPNSP--VVPPTPVKNPVKKWKPPWEDDDA
                                                                                                                                                                                                                                         TPKEPAPTTTKKPAPTA-PKEPAPTTPKETAPTTPKKLTPTTP-----EKLAPTTPEKPA
                                                                                                                                                                                                                                                                            PEKK-TPVLAKKAPTKPDSEAAADPVSGPSSKDPKLAKKAPVKPRDPSPMKAVPIKPAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TPVKNPVKKWRPPWEDDETPADDVSKPTDAKKTPSLAKKDPAPAKESLKPKADTKAPAKP
                                                                  PSPKKAEPNSPVVP----PTPVKNPVKKW-KPPWEDDDEPTEEVKKPSE--PEKKTPVLA
                                                                                                                                                           PTTPEELAPTTPEEP----TPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTT-PKE 496
                                                                                                                                                                                                                                                                                                           TTKEPAPTTTKSAPTTP-KEPSPTTTKEPAPTTPK--EPAPTTPKKPAP--TTPKEPAPT
                                                                                                                                                                                                                                                                                                                                               APAAVKKPEPISKPKDTAPKKAEPNSPVVP-PTPVKNPVKKWKPPWEDDDAPAKPVSLPE
                                                                                                                                                                                                                                                                                                                                                                     APTAPKKPAP-----TTPKEPAPTTPKEPAPTTTKEP----SPTTPKEPAPTTTKSAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTTKEPA-----TTKSAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EAPAKKTPVLKRKEPAAKDTAKPATSKTPETPEKKDPVKPRDSSPKKVAAKPDSAQAPA-
                                                                                                                                                                                                                                                                                                                                                                                                                   RDPSPKKVAPTAPEKKTPVLAKKEPAGPADSKTKEPEKSKPRDPSPKKAVPAKPVPKTEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KEPTPTTP----KEPASTTPKEP----TP-TTIKSAPTTPKEPAPTTTKSAPTTPKEPAP 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ETQSIEKTSAKDLAPTSKYLAKPTPKAE----TTTKGPA------LTTP 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPVVPPTPVKNPVKKYKPPWEVDDEPAEEVKKPSAPEKKTPVLKRKEPEPSSTTPSSDPS 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PKITTAKPI-NPRPSLPPNSDTSKETSLTVNKETTVETKETTTNKQTSTDGKEKTTSAK 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KNPTKKWKPPWEDETPVEEVKEPPVPEKKAPVLKKKDPAPAAKARDPSPSKAAPKKVEPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     KNRTKK-KP----TP-----KPPVVDEAGSGLDNGD----FKVTTPDTSTTQHNKVSTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            from GB/EMBL/DDB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
 974
                                     603
                                                                    918
                                                                                                                                       865
                                                                                                       555
                                                                                                                                                                                                                                            442
                                                                                                                                                                                                                                                                            750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              512
                                                                                                                                                                                                         805
                                                                                                                                                                                                                                                                                                               388
                                                                                                                                                                                                                                                                                                                                                  691
                                                                                                                                                                                                                                                                                                                                                                                  333
                                                                                                                                                                                                                                                                                                                                                                                                                                                      283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56,
```

```
Ъ
                                                                                                                                                           ρy
                                                                                           Ωy
                                                                                                                                밁
                                                                                                                                                                                                B
                                                                                                                                                                                                                                                                Ъ
                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R:Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A. submitted to the EMBL Data Library, June 1994 A;Description: Pex genes; pollen-specific genes with extensin-like A;Reference number: $49915 A;Accession: $49915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: EMBL:234465; NID:g600117; PIDN:CAA84230.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C:Species: Zea mays (maize)
G:Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
C:Accession: S49915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
S49915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-1188 < RUB>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               extensin-like protein - maize
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 29.9
'-hog 230; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Accession: S49915
                           352
                                                            719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1194 -LNKKAPAEKPTEK------PKPKEVSKEPPKPTEPPKP-AAPKKWKPPWEDDPDEPE 1243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1090
                                                                                                                                                               241
                                                                                                                                                                                                                                 184
                                                                                                                                                                                                                                                                                                    129
                                                                                                                                                                                                                                                                                                                                     511
                                                                                                                                                                                                                                                                                                                                                                                                       460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1244 ADFTMPAPKKPDTEDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1032 AAK-PDPKIPEV-PPTPVKNPVKKWKPPWEDDDEPSEPVSAPEPEKKTPVLAKKAPTKPA 1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                798
                                                                                                                                                                                                                                                                                                                                                                      69
                                                                                                                                                                                                                                                                                                                                                                                                                                          13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               604
EPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTTKKPAPTAPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          u
                                                 SP--PKEPVSSPPQTP----KSSPPPAPVSSPPPTPVSSPPALAPVSSPPSVKSSP----
                                                                                APTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPK 351
                                                                                                                                                  PKKPAPTTPKEPAPTT-----PKE---PTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKP 291
                                                                                                                                                                                                                                                                                                                                                        SDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKV 126
                                                                                                                                                                                                                                                                                                                                                                                                                           PTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPSL----PPN
                                                                                                                 PAKSTPPPEEYPTPPTSVKSSPPPEKSLPPPTLIPSPPPQEKPTPPSTPSKP-PSSPEKP
                                                                                                                                                                                     PPAPVASPPPPVKSPPPPTPVASPPPPAPVASSPPPMKSPPPPTPVSSPPPPEKSPPPPP 659
                                                                                                                                                                                                                    PTTTKSAPTTP--KEPAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTP-KEPAPTT 240
                                                                                                                                                                                                                                                       SPPPPVKSPPPPAPVGSPPPPEKSPPPPAPVASPPPPVKSP--PPPTLVASPPPPVKSPP 599
                                                                                                                                                                                                                                                                                      ----LAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPA 183
                                                                                                                                                                                                                                                                                                                                                                                          PTPHSPPAD------DYVPPTPPVPGKSPPATSPSPQVQPPAASTPPPSLVKLSPPQ 510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KWKPVWDDDPDEPEADFTVPAPSKKPDTEDPADPLG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---TTPTPRKMTSTMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE 797
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IMNKPEETAKPKORATNSKATTPKPOKPTKAPKKPTSTKKPKTMPRVRKPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TKPDSEAAADPVSGPTSKDPKLSKKAPVEKPKPTTDPKDDKLKPSPAKKPEKAPEPAAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            V------PTTKTPAAT-----KPEMTTTAKDKTTERDLRTTPETT--TAAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PPTPVKNPVKKWKPPWEDDDEPAEPVSAPEPEKKTPVLAKKAPAKPRDP----SPKKAAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APTTTKGPTS-----TTSDKP-----APTTPKET-----APTTPKEPAPTTPKKPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                15.6%; Score 784; DB 2; 29.9%; Pred. No. 1.9e-26; tive 59; Mismatches 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          861
                                                                                                                                                                                                                                                                                                                          -SPPPPVKTTSPPAPIG----SPSPPPPVSVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 1188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GPKTKDPK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  122;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PID:g600118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           domains
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                 768
                                                                                                                                                                                                                                                                                                                          541
                                                                                                                                                                                                                                                                                                                                                                                                                               68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            737
                                                                                                                                                                                                                                                                                                                                                                                                                                                              33,
```

Db Qy	Qu Be Ma	C; Datt C; Acco R; Goes Submi A; Ref A; Acco A; Sta A; Moll A; Res A; Cro C; Gen A; Gen	Db RESU RESU T311	Фр	Qу Db	Qу	Оу	Оу	Оу	Qу	Оy	ОУ	Qу	рb
9 TKKKPTPKPPVVD : 338 TPYAPTEKPYDVE	uery Match est Local Similarity atches 355; Conserva	ce: 22-Oct-1999 cession: T31108 arnhardt, B. itted to the EMB ference number: cession: T31108 atus: preliminar lecule type: DNA sidues: 1-1489 < coss -references: netics: netics: netics:	1425 1425 19 germi	26 73	767 QDTTPFKITTLKTTTLAP : 1338 DEPTPSDEPT	707 ATKPEMTTTAKD	647 TPETPPPTTSEV : : 1230 DTPSDEPTPSD-	592 PT-TPKKPAPKE : :: 1176 PSETPEEP	533 ELAPTTTKEPT-: : :	476 NTPKEPAPTTPK	417 TAPTTPKKLTPTTPEKI	358 TKEPAPTTPKEP/ : : 969 SDEPTPSDEP/	302 TTPKEPAPTTT: : : : : : : :	865 SDEPTPSDEPTPS
KKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTT- 	17.0%; Score 853; Di _y 32.3%; Pred. No. 2.90 ervative 46; Mismatches	_revision 22 brary, April ated from GB 1185; NID:g3	EAMLQTITERWLINGS FROM THE CONTROL OF	KAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTS	TTTLAPKVTTTKKTITTTEIMNKI	DKTTERDLRTTPETTTAAPKMTKETATTT	STPTTTKEPTTIHKSPDESTPELS	-TPKKPAPKELAPTTTKGPTSTTSDKPAPTT :: :: ::	APTTTKEPT-STISDKPAPITPKGIAPTIPKEPAPTIPKEPAPTIEKGTAPTILKE 	PKEDAPTTPKEPAPTTPKEPAPT-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAP 	PEKLAPTTPEKPAPTTPEELAPT :: :: : SDEPTPSDEPTPSDEPT	TKEDAPTTPKEPAPTTPKKPAPTTPKEPAPT-TPKE	PTTTKEPSPT-TPKEPAPTTTKS-APTTTKEPAPT : : : : : :	SDEPTPSDEPTP-SDEPTPS
QHNKVSTSPKITTAKPINP : : ERMHYAHIEKPCDTEVTMYAPTEE	3 2; Length 1489; e-29; 475; Indels 222;	999 #text_change 22-0c DDBJ ; PID:g3851514; PIDN:A	1 th	SGGGGGGGGTVPTSPTPTPTS	AKPKDRATN PTPSDEPTP	ATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTT	TPETPEPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPA 	TPKETAPTTPKEPAPT-TPKKPAPT : : :: EPTPSD-EPTPSDEPTPSETPEEPIPT	PTTPKEPAPTTPKGTAPTTLKEPA : : PSDEPTPSDEPTPSDEPT	TPKGTAPTTLKEPAPTTPKKPAPK	TAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPT-TPEEPAPTTPKAAAP	PAPT-TPKEPAPTTTKKPAPTAPKEPAPTTPKE 	APTTTKEPAPTTTKSAPTTPKEPSPTT	סי
R 62 7 397	аp	E-1999 AC72308.1	65	1424	1 82	766	1277	646 1229	591 1175	532 1122	475 1064	416 1019	357 968	⊢

```
망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ρy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ΩV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                Qy
                                                                                                                                                                                                     DЬ
                                                                                                                                                                                                                                 γQ
                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                         Ър
                                                                                                                                                                                                                                                                                                                                                                     Qy
                                                                                                                                                                                                                                                                                                                                                                                                       Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                В
                                                                                                20
                                                                                                                                 Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     574 TYAPTEETT-YAPTEETMYAPIEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYA 632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       398 TYAPTEETTYAPTEETTYAPTEETPYEPTEETTYTPTEETTYAPTEETTYAPTEKTTYAP
                                                                                                                                                                                                                                                                                                                                                                                                            1051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101 GKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPA 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     633
 1403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                931
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              449 LAPT--TPEEPTPTTPEEP-----APT-----TPKAAAPNTPKEPAPTTPKEP- 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                813 TPYEPTEETTYAPTEETPYEPTEET--TYTPTEETTYAPTEETTYAPTEKTTYAPTEETT 870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  693 TTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETMYAPIEETTYGPTEETTYA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STTPKEPTPTTIKSAPTTPKEPAPT-TTKSAPTTPKEPAPT--TTKEPAPTTPKEPAPTT 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-----APTAPKEPAPTTPKEPAPTTPKEP-----APTAPKKPAPTTPKEPAPTTP 304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TKEPAPTTTKSAPT------TPKEPAPTTPKKPAPTTPKEPAPTTPKE 267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EETPYEPTEET-TYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPAEETPYEPTEET 573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TEETTYAPTEETPYEPTEETTYAPTKETTYAPT - - - EETTYASTEETTYAPTEETTYAPA 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TPKEPAPITIKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEE 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTEATTYAPTEETPYAPTEETTYEPTGETTYAPTEETTYAPTEETTYAPTEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTEETTYASTEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEE 692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PSLP---
                                                                                                                                                                                                                                                                                                                                                                                                          EETTYAPIEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETTYAPTEETTYAPTEETT 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VAPTEETMYAPIEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETTYAPTEETTYAST 1050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EETTYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPAEETPYEPTEETTYAPTEETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YAPTEETPYEPTEETTYAPTKETTYAPTEETTYASTEETTYAPTEETTYAPAEETPYEPT 930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PT-TTKSAPT--TPKEPSPTTTKEP------APTTPKEPAPTTPKKPAPTTPKEPAPT--
                                                                                                                                                                                                                                                                                                     AEPTPKA-LENSPKEPG-----VPTTKTPAATKPEMTTTAKDKTTERDLRTTP----ETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEP-----APTTPKGTAPT------
                                                                                                                                                                                                                                                                         TEETPYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPTEETTYAPTEET
                                                                                                                                                                                                                                                                                                                                         YAPAEETPYEPTEETTYAPTEETTYAPTEETMYAPIEETTYGPTEETTY-APTEATTYAP 1169
                                                                                                                                                                                                                                                                                                                                                                     KETAPTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTKEPT--TIHKSPDESTPELS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----APTTPKEPAPTTPKETAPTTPKGTAPT--TLKEPAPTTPKKPAPK-ELAPT--TT
                                                                                                                                                                                                                                   TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTT-----PFKITT---LKT 779
QGINGIGVENKVRYNNAG
                              OTPNSKLVEVNPKSEDAG
                                                                 PARESTSTVSTEKPCNTEEFTDEPTDEPTDE--PSDEPTDEPTDEPTDLPTDEPSTPCDN 1402
                                                                                                                                   TTYAPTEATTYAPTEETPYAPTEETTYEPTGETTYAPTEETTYAPTEETTYAPMEETPYE 1344
                                                                                                                                                                   TTLAPKVTTTKKTITTTEIMNKPEETAKPKDRAT--NSKATTPKP-QKPTKAPKKPT--- 833
                                                                                                                                                                                                     TYAP----TEETTYEPTEETTYAPTEETTYAPTEETTYAPTEETMYAPIDETTYGPTEE 1284
                                                                                                 -STKKP-KTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRP--N 881
                                                                                                                                                                                                                                                                                                                                                                                                                                          -TLKEPAPTTPKKP-----APKE----LAPT--TTKGPTSTTSDKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --PNSDT----SKETSLTVNKETTVETKETTT---TNKQTSTD 100
     1420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            388
                                                                                                                                                                                                                                                                                                                                                                                     682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         990
                                                                                                                                                                                                                                                                                 1229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         585
```

```
밁
                                                                                               Š
                                                                                                                                                                               B
                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Accession: PQ0329
A;McCession: PQ0329
A;McCession: PQ0329
A;McCession: PQ0329
A;McGesidues: 2328-2342,'K',2344-2354 <XUG1>
C;Genetics: A;Gene: GDB:MC2
A;Gene: GDB:MC2
A;Gene: GDB:HDC2
A;McD position: 11p15.5-11p15.5
C;Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; von C;Superfamily: von Willebrand factor type C;Superfamily: von Willebrand factor type C repeat homology <VWC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A.Experimental source: bronchus R;Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang, D.; Jones, C.; Forstner, Blochem. Blochem. Commun. 183, 821-828, 1992
A;Title: Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-t A;Reference number: PQ0328; MUID:92198477
A;Accession: PQ0328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Molecule type: mRNA
A;Residues: 2328-2468 <XUG>
A;Cross-references: GB:M86523
A;Experimental________ small intestine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Molecule type: mRNA
A;Residues: 'T',1925-1948,'TTS',1952-1954 <JAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      J. Clin. Invest. 87, 77-82, 1991
A;Title: Human bronchus and intestine express the same mucin gene. A;Reference number: A61257; MUID:91086481
A;Accession: A61257
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 21.2
Best Local Similarity 29.2
Matches 342; Conservative
                               1760
                                                                                                                                                                                                                                                                                                                                  1649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1472 P-PTTTPSPPTTTPSPPT-TTTTTPPPTTTPSPPMTTPITPPASTTTLPPTTTPSPPTTT 1529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1412 TTTTLPPTTTPSPPTTTTTPPPTTTPSPPITTTTPLPTTTPSPPISTTTPPPPTTTPS 1471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1352 GOKVQCDVSVGFICKNEDQFGNGPFGLCYDYKIRVNCCWPMDKCITTPSPPTTTPSPPPT 1411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1292
                                                                                   494 PKEPAPTTP----KETAPTTP------KG----
                                                                                                                                                                                                                                                          434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1240 WEICGPNGTVEKHENICSITTRPSTLTTFTTITLPTTP-----TSFTTTTTTTPTS 1291
                                                                                                                                                                                                                                                                                                                                                                                                         374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          138 TTTKGPALTTPKEPTTTTPKEPASTTPKEPTPTTIKSAP-TTPKEPAPTTTKSAPTTPKE 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32 FKVTTPDTSTTQH---NKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETK 88
PLPPSITPPTFSPFSTTTPTTPCVPLCNWTGWLDSGKPNFHKPGGDTELIGDVCGPGWAA 1819
                                                                                                                                                                                                                                                                                                  TIPPPTITPSPPTTTTPSPPTTTTPPPTTTPSSPITTTP--SPPTT----TMTTPS-- 1700
                                                                                                                                                       PTTTPSSPITTTTPSSTTTP-SPPPTTMTTPSPTTTMTTLPPTTTSSPLTTT
                                                                                                                                                                                                                      APTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTT 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTITTTPPPTTTPSPPTTTTPPPTTTPSPPTTTPITPPTSTTTLPPTTTPSPPPTTT 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SPTTPKEPAPTTTKSAP-TTTKEPAPTTTKSAP-TTPKEPAPTT 373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ТРКЕРТРТТРКЕРАРТТКЕРАРТТРКЕРАРТАРККРАРТТРКЕРАРТТРКЕРАРТТТКЕР 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAPTTTKEPAPTTPKEPAPTTTKEPAPTTTKSAP-TTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                     PKKPAPTTPKEPAPTTPKEPAPTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKL 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TTTPPPTTTPSPPTTTPTTP-PTSTTTLPPTTTTPSPPTTTTPPPTTTTPSPPTTTTPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STVLSTTPKLCCLWSDWINEDHPSSGSDDGDREPFDGVCGAPEDIECRSVKDPHLSLEQH 1351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21.2%; Score 1004...,
29.2%; Pred. No. 8e-38;
1tive 74; Mismatches 442; Indels 315; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------LAPTSKVLAKPTPKAE 137
                                                                                                                                                       1759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   39;
```

R; Fu J. E A; Ti e. A; Re A; Ac A; St A; Re
R;Fujino, T.; Beguin, P.; Aubert, J.P. J. Bacteriol. 175, 1891-1899, 1993 A;Title: Organization of a Clostridium tle. e. A;Reference number: 218847; MUID:9320993; A;Accession: T18262 A;Status: preliminary; translated from Gi A;Molecule type: DNA A;Residues: 1-1664 <fuid< td=""></fuid<>

```
OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
                                                                                                                                                          762.5
671
666.5
639.5
632
632
622.5
622.5
627.5
607.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
556.5
556.5
556.5
555.5
544
545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           April 26, 2002, 16:13:30 ; Search time 82.85 Seconds (without alignments) 865.181 Million cell updates/s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
21.2
178.0
179.0
155.6
155.6
155.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
115.2
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     219241 seqs, 76174552 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-556-246-1_COPY_200_1140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VKDNKKNRTKKKPTPKPPVV......DMDYLPRVPNQGIIINPMLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                      1367
1129
1489
17962
7489
7662
3507
8390
6642
35970
801
35970
801
35970
801
35970
801
35970
801
35970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1188
2187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3020
1664
1489
1274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
    A43932
T11262
T31108
T31108
T316251
S48478
T30826
T30826
T30826
T348478
T11622
C84672
C84672
T297518
T34518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ħ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            219241
                                                                                                                                                                                                                                                                                                                                                                                                       mucin 2 precursor.
S-layer protein -
cyst germination;
hypothetical protein -
extensin-like protein -
mucin |
high molecular manucin |
mucin |
high procursor
  hypothetical prot
nucleolar phospho
hypothetical prot
protein UNC-89 -
hypothetical prot
mucin MUC5B, trac
mucin MUC5B, trac
qene LF3 protein
larval glue prote
hydroxyproline-ri
hypothetical prot
polyphenolic adhe
ascites sialoglyc
hydroxyproline-ri
hypothetical prot
polyphenolic adhe
ascites sialoglyc
hydroxyproline-ri
hypothetical prot
calphotin - fruit
                                                                                                                                                                                                                                                                                                                           hypothetical prot
extensin class 1
hypothetical prot
elastic titin - h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   941
```

44 45	43	42	41	40	υ 9	ж 8	37	ω 6	u U	34	33	32	31	30
504.5	505	509	509.5	509.5	512	518	518	522	530	530.5	532	533	538.5	542
10.1	10.1	10.2	10.2	10.2	10.2	10.3	10.3	10.4	10.6	10.6	10.6	10.6	10.7	10.8
606 1118	700	1459	867	662	866	1072	813	328	865	1162	756	1087	416	369
NK	N	2	N	Ν	N	Ц	Ν	N	Ν	N	2	ш	2	2
A48292	A54641	T32271	T45463	A45155	T45462	A37221	S70795	JQ0985	A47282	JH0557	T27642	QFMSH	JU0465	s20500
mucin, tracheobron	interspersed repea	hypothetical prote	membrane glycopiuc	mucin Fim-C.i At	membrane grycoproc	neuroniliament tip	VSAA PIOCETII PIECU	Hydroxyprorine ric	Calcinication by	exC-aipia starrans	myportierrear proce	detail cut the design of the d	extensin precureor	hydroxyproline-ric

or,	s ec
RESULT 1 RESULT 1 Ad3912 RESULT 1 CSPACEUS: Homo Sabless (man) A. Fills: Molecular Cloning of human intestinal mucin (MUC2) CDNA. Identification of t A. Fills: Molecular type: man A. Residues: 16.03 CGPACIAGO (MAN) A. Residues: 16.03 A. CCCC (MAN) A. Residues: 16.03 A. CCCC (MAN) A. Residues: 10.04 A. Residues: 10.05 A. Residues: 10.07 A. Residues:	36 522 10.4 328 2 JOV995 37 518 10.3 813 2 S70795 38 518 10.3 1072 1 A37221 neurofilament trip 38 518 10.2 866 2 T45462 mucin FINC.1 - Af 40 509.5 10.2 867 2 T45463 mucin FINC.1 - Af 41 509.5 10.2 867 2 T45463 mucin FINC.1 - Af 42 509 10.2 1459 2 T32271 hypothetical prote 43 505 10.1 700 2 A54641 interspersed repea 44 504.5 10.1 606 2 A43427 neurofilament trip 45 503.5 10.0 1118 2 A48292 mucin, tracheobron

٠,

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
US-08-276-967-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2, Application US/08276967

Patent No. 5851817

GENERAL INFORMATION:
APPLICANT: Hardy, Daniel M.
APPLICANT: Garbers, David L.
TITLE OF INVENTION: Species-Specific Egg-Binding Proteins
TITLE OF INVENTION: Sperm
                 ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSD:418\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US,
                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            869
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 889
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               589
                                                                                                                                                                                                                                                                                                                                                           STREET:
CITY: 1
                                                                                                                               CLASSIFICATION:
                                                                                                                                                                                                      COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0
                                                                                                                                                FILING DATE: Submitted Herewith
                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APDXRPXPG-----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VTSAPDXRPXPGSTAPXAHGVTSA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPID 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KPAP--TTPK----ETAPTTPKEPAPTTP------KKPAPTTPETPPPTTSEVSTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEPAPTTPKG----TAPTTLKEPAPTTP-----KKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APXAHGVTSAPDXRPXPGSTAPXAH-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTKEPTT-----IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGS--TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS----APDXR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTP-----KKPAPKELAPTTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPX------PGST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -TKPEMTTTAKDK----TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                                                                                                                                                                     77210-4433
                                                                                                                                                                                                                                                                                                                                                           Houston
                                                                                                                                                                                                                                                                                                                                          Texas
                                                                                                                                                                                                                                                                                                                                                                             E: Arnold, Whi
                                                                                                                                                                                                                                                                                                                        USA
713-789-2679
                                                                                                                                                                                                        PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                            White & Durkee
                                                                                                                                                                      us/08/276,967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---GVTSAPDXRPXPGSTA----PXAHG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -PDXRPXPGSTAPXAHGVTSAPD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     644
```

```
ΩV
                                                                                 밁
                                                                                                                       Qy
                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                          Ъ
                                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-276-967-2
  밁
                                                                                                                                                                   Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2476 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 6.2%;
Best Local Similarity 32.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         471 PKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEP
  694
                                       932 SRIAEAMLQTTTRPNQTPNSKL 953
                                                                                                                                                                                                          812 EKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETA 871
                                                                                                                                                                                                                                                           572
                                                                                                                                                                                                                                                                                                                                          513 VPTEKPIVPTEKHTIPTEKLTVLTERTTTPTERTTIPTEKPTVPTEKPSVPT-EKPTVPT
                                                                                                                                                                                                                                                                                                                                                                                    702
                                                                                                                                                                                                                                                                                                                                                                                                                              459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             366 TIPTEKSTVPT--KKPTVFKEPTLPPE-GPTVPAE-RPTTPPEGPAVPPKG--PTVLTE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      531 APTTPKAAAPNTPKEPA----PTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  313 PSETSVSTEKPVAPTE----KPTVPSEIYTIPTEKPMVHMEKPIVHT--EKPTVPT-EKP
AHFERCACPVSCQ-SPTPNCEL
                                                                                                                         KPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPT
                                                                                                                                                                   ERTT-----TPTIRTTTPTERTT---IPTKKTT-----VPTEKTIIPT-----ERTI
                                                                                                                                                                                                                                                                                                PKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTT 811
                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTAPTTLKEPAPTTPKKPAPKELAPT-TTKGPTSTT---SDKPAPTTPKETAPTTPKEPA
                                                                                                                                                                                                                                                                                                                                                                                  PTTPKKPAPTTPETPPP-----TTSEVSTPT----TTKEPTTIHKSPDESTPELSAEPT
                                                                                                                                                                                                                                                                                                                                                                                                                            -TERTTIPAEKPTV---PIEKPMVPTERTTIPTERTTIPTEKPTVPTEKLTVPT--EKPI 512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPT-TPKEPAPTTPKEPAPTTPK 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    163;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                         -EEPTIPTEKLTVPT--ERTTTPTKRTTTPTIRTTPTI-----RTTTPT
                                                                                 TTPQP-SPTLVPTQPAAVVMPST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---WPTSHTEKSTVHTEKPILPTGKSTIPTEKPMVPTKRT-----TTP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Score 419.5; DB 2;
; Pred. No. 1.4e-19;
43; Mismatches 177;
                                                                                     ---SATTVTPRTTIASCP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 2476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels 119;
```

Gaps

31;

365

Search completed: April Job time: 489 sec 26, 2002, 16:25:49

-PN

613

571 751 701 458

```
밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
US-09-083-116-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 2 Patent No.
                                   TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENERAL INFORMATION:
                                                                                               NAME: Teskin, Robin L.
REGISTATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEPHONE: (703) 836-6620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
                                                                                                                                                                                                                                                                                                                                                    PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PO
                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: LATHE, Richard
APPLICANT: HAREUVENI, Mara
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION FOR THE
TITLE OF INVENTION: TREATMENT OR PREVENTION OF A MALIGNANT TUMOR
NUMBER OF SEQUENCES: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1008 ETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPID 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   958 VTSAPDXRPXPGSTAPXAHGVTSA------PDXRPXPGSTAPXAHGVTSAPD 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     912 APXAHGVTSAPDXRPXPGSTAPXAH------GVTSAPDXRPXPGSTA----PXAHG 957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             949 -PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD 1007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            889 KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT 948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          763 DXRPXPGSTAPXAHGVTSAPDXRPXPGST--APXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US/09/083,116 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET: P.O. Box
CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                      ENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -GSTAPXAHGVTSAPDXRP--XPGSTAPXAHGVTSA-PDXRPX------PGST 911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APDXRPXPG-----STAPXAHGVT-----SAPDXRPXPGSTAPXAHGVTSAPDXRPXP- 868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQ 888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -TKPEMTTTAKDK----TTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    , Application US/09083116
6203795
: 2035 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KIENY, Marie-Paule
LATHE, Richard
HAREUVENI, Mara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAMBON, Pierre
                                                                                                                                                                                                                                                                                                                                                                             WO PCT/FR91/00835
                                                                                                                                                                                                                                                                                                                                                                                                                                           08/479,537
                                                                                                                                           017753-025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             820
```

88	34 TPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP 5	Qу 5
71	20 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGST 4	Db 4
ຜິ	8 TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT 5	Qy 47
177	6 PAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPT (Qy 42 Db 36
425 359	7 TKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKE	
376 300	PKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTT :	Оу 3 Db 2
329 240	283 TKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPT 2	Оу 2 Db 1
282	242 TTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTT 2	Qy 2
241 122	9 P-KAETITKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAP	Qy 19 Db 6
55 .98	140 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 1	Qy :
s 53;	Query Match 6.2%; Score 424.5; DB 4; Length 2035; Best Local Similarity 23.7%; Pred. No. 5.3e-20; Matches 255; Conservative 69; Mismatches 511; Indels 241; Gaps	Query Best Matcl
	NAME/KEY: Peptide LOCATION: 1.21 COTHER INFORMATION: /note= "Amino acids 1 to 21 are a - OTHER INFORMATION: 21 amino acid precursor sequence." 083-116-2	s-09-
• CCT, CCC,	NAME/KEY: Peptide LOCATION: 17 COTHER INFORMATION: /note= "Amino acid 147 is X2 = Xaa OTHER INFORMATION: which is the codon for Pro or Ala wherein Pro OTHER INFORMATION: or CCG; and Ala = GCT, GCC, GCA, or GCG." EATURE:	'zş
· - ACT, ACC,	Peptide 144 RMATION: /note= "Amino acid 144 is Y RMATION: which is the codon for Thr o RMATION: or ACG; and Asn = AAT or AAC	
erein Pro - C	EY: Peptide ON: 134 INFORMATION: INFORMATION: INFORMATION:	
nerein the repe	Y: Peptide N: 128.1899 NFORMATION: /note= "The amino acids spanning NFORMATION: 128 to 1899 constitute a repeated region w) NFORMATION: 20 amino acids, 17 of which are fixed. The NFORMATION: repeats varies from 1 to 40."	······································
	STRANDEDNESS: single TOPOLOGY: linear MOLECULE TYPE: peptide FEATURE:	

•

```
멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
US-08-479-537A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 2, Application US/08479537A Patent No. 5861381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                              TELEFAX: (703) 836-202
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-OCT-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION TITLE OF INVENTION: TREATMENT OR PREVENTION OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 04-APR-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE: 23-OCT-
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQUENCES:
                                                                                                                                                   NAME/KEY: Peptide
LOCATION: 128.1899
OTHER INFORMATION: /n
OTHER INFORMATION: 12
OTHER INFORMATION: 20
                                                                                                                                                                                                                                                                                                                                                                                                    TELECOMMUNICATION INFORMATION: TELEPHONE: (703) 836-6620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRIOR APPLICATION DATA:
                  FEATURE:
                                                                                                                        FEATURE:
                                                                                                                                                                                                                                                             MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: FR 90 FILING DATE: 23-OCT-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: FILING DATE: 04-AP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US/08/479,537A FILING DATE: 07-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CITY: Alexandria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: Teskin, Robin L. REGISTRATION NUMBER: 35,030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: FILING DATE: 14-MAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE:
                                                   LOCATION: 134
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                            TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE/DOCKET NUMBER:
                               OTHER INFORMATION:
                                                                                                    NAME/KEY:
                                                                                                                                        OTHER INFORMATION:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VTSAPDXRPXPGSTAPXAHGVTSA--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22313-1404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P.O.
                                                                                                                                                                                                                                                                                                                             2035 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HAREUVENI, Mara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LATHE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KIENY, Marie-Paule
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAMBON, Pierre
                                                                                                                                                                                                                                                                          linear
                                                                                                      Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BURNS,
Peptide
                                                                                                                                                                                                                                                                                                                                                                                                   (703)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IUMBER: US 08/403,576
14-MAR-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Box 1404
                                                                                                                                                                                                                                                             peptide
                                                                                                                                                                                                                                                                                            single
                                                                                                                                                                                                                                                                                                                                                              3) 836-2021
1836-2021
190: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Richard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOANE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US 08/039,320
                                  /note= "Amino acid 134 is X1 = Xaa
Xaa Xaa which is the codon for Pro or
CCC, CCA, or CCG; and Ala = GCT, GCC,
                                                                                                                                        /note= "The amino acids spanning 128 to 1899 constitute a repeated region wherein the 20 amino acids, 17 of which are fixed. The number of repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FR 90/13101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SWECKER & MATHIS, L.L.P
                                                                                                                                                                                                                                                                                                                                                                                                                                         017753-025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Version
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -PDXRPXPGSTAPXAHGVTSAPD 1003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FOR THE A MALIGNANT TUMOR
                                      Ala wherein
GCA, or GCG
                                               Pro =
                                                                                                                                                              repeat
such
                                                          CCT,
```

```
Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   뫄
                                                                                                                                                                                                                                                                                                                                                                                                                                       QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-08-479-537A-2
                          Qy
                                                                                   Qγ
                                                                                                                Вb
                                                                                                                                           δÃ
                                                                                                                                                                         밁
                                                                                                                                                                                                    ρy
                                                                                                                                                                                                                                   Вр
                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                            QY
                                                                                                                                                                                                                                                                                                                                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
맑
                                                          Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 255; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: Peptide
LOCATION: 1..21
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             199 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     140 TSLTVNKETTVETKETTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT
                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                           478
                                                                                                                                                                                                                                                                                                                                                                                  426
                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                             377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        123 --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                        589
                                                                                                                                                                                                                                    472
                                                                                                                                                                                                                                                                534
                                                                                                                                                                                                                                                                                                                                                        360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                   586
                                                                                                                                                                             531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16 TVLTV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER
 703
                             727
                                                                                        683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OTHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OTHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY:
                                                                                                                                                                                                                                                              TPKA----AAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGT-APTTLKEPAP
                                                                                                                                                                                                                                                                                            SAPDXRPXPGSTAPXAHGVTS----APDXRPXPGSTAPXAHGVTSAPDXRPX----PGST
                                                                                                                                                                                                                                                                                                                                                                                 PAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPK----EPAPTTTKKPAPTAPKEPAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 \tt PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP
                                                                                                                                                                                                                                   APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS
                                                                                                                                                                                                                                                                                                                        TPKETAP----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPT
                                                                                                                                                                                                                                                                                                                                                   SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT
                                                                                                                                                                                                                                                                                                                                                                                                              {\tt APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT}
                                                                                                                                                                                                                                                                                                                                                                                                                                          TKEPSPTTPKEPAPTT --TKSAPTTTKEPAPTT----TKSAPTTPKEPSPTTTK----E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKE--PAPTTKEP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTPKEPAPTTT - - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS
                                                                                                                PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX 644
                                                                                                                                                                           TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTS---
                                                                                                                                                                                                          TTP-----
{\tt DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP}
                                                                                      KPAP -- TTPK ---- ETAPTTPKEPAPTTP ---
                                                                                                                                              KEPAPTTPKG----TAPTTLKEPAPTTP--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R INFORMATION:
R INFORMATION:
R INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide
147
                                                                                                                                                                                                     -KKPAPKELAPTTKEPTSTTSDKPAP--TTPKGTAPTTPKEPAPTTP 636
                              ---IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -VTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   6.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---KSAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "Amino acid 144 is Y = Xaa which is the codon for Thr or Asn or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Amino
21 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "Amino acid 147 is X2 = Xaa which is the codon for Pro or Ala wherein or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -APTTPKEPAPTAPK----KPAPTTPKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 424.5; DB 2;
Pred. No. 5.3e-20;
9; Mismatches 511;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             acids 1 to
                                                                                                                                                  -KKPAPKELAPTTTKGPTSTTSD
                                                                                           -KKPAPTTPETPPPTTSEVSTP 726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -APTTTKEPAPTTPKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             o 21 are a sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                               --TAPXAHGVTSAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            wherein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2035;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241;
                                                                                                                                                                                 -APDXR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Thr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pro =
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                     682
                                                                                                                                                                                                                                         530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              198
                                                                                                                                                                                                                                                                                                                               533
                                                                                                                                                                                                                                                                                                                                                                                        477
                                                                                                                                                                                                                                                                                                                                                                                                                                                425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                 585
                                                                                                                                                                                                                                                                                                    471
                                                                                                                                                                                                                                                                                                                                                             419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    329
                                     772
                                                               702
                                                                                                                                                                                                                                                                      588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACC
```

```
; NAME/KEY: Peptide
; LOCATION: 1..21
; OTHER INFORMATION:
; OTHER INFORMATION:
US-09-083-116-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 0177:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEPHONE: (703) 836-6201
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 amino acids
TYPE: amino acids
                                                                                                                    FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Peptide
LOCATION: 128.1727
OTHER INFORMATION: 17
OTHER INFORMATION: 17
OTHER INFORMATION: 20
OTHER INFORMATION: 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US 08/039,320
FILING DATE: 04-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,576
FILING DATE: 14-MAR-1995
ATTORNEY_AGENT INFORMATION:
NAME: Teskin, Robin L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MOLECULE TYPE: peptide
                                                                                                                                                                                                     NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
GOVERNATING STEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                       NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR91/00835
FILING DATE: 23-CCT-1991
PRIOR APPLICATION DATA:
PRIOR APPLICATION THE OB //120 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
                                                                                                                                OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                  OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                  OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/09/083,116 FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: P.O. Box
CITY: Alexandria
STATE: Virginia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P.O. Box 1404
                                                                                                                                                                                                       Peptide
147
                                                                                                                                                                                                                                                                                                                                       Peptide
144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          United States
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide
134
         /note= "Amino acids 1 to 21 are a 21 amino acid precursor sequence."
                                                                                                                 /note= "Amino acid 147 is X2 = Xaa which is the codon for Pro or Ala wherein Pro = or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                    /note= "Amino acid 144 is \gamma = xaa which is the codon for Thr or Asn wherein Thr = ACT, ACC, or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                    /note= "Amino acid 134 is x1 = xaa
which is the codon for Pro or Ala wherein Pro = CCT, CCC,
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note- "The amino acids spanning 128 to 1727 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08/479,537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     017753-025
                                                                                                                                            CCT, CCC,
                                                                                                                                              ဌ
                                                                                                                                                                                                                                                                              AC.
                                                                                                                                                                                                                                                                                                                                                                                                              S
```

065	912 APXAHGVTSAPDXRPXPGSTAPXAH	рь 9 Оу 10	
ίLSD	PRVPNQGIIINP	Оу 9	_
POST	889 KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT	Оу 8 Db 8	
ΧP-	29 21	Оу 8 Db 8	
VTS	63		
Αp	APXAHG		
A - 4		Db 6 Qy 7	
מל	86 PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDX	оу , 5 Оу , 5	
83	637 KEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSD		
3 1	31 TAPXAHG		
ୁ ନ	72 APXAHGVTSAPE		
S_ 7	420 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGST	O D	
9 ≶	360 SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 478 TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPEAPTTPEENAPTTPEEPEAPTTPEENAPTTPEEPEAPTTPEENAPTTPEENAPTTPEENAPTTPEEPEAPTTPEENAPTTPEEPEAPTTPEENAPTTPEEPEAPTTPEENAPTTPEEPEAPTTPEEPEAPTTPEENAPTTPEEPEAPATTPEEPEAPTTPEEPEAPATTPEEPEAPAPTTPEEPEAPAPTTPEEPEAPATTPEEPEAPATTPEEPEAPATTPEEPEAPATTPEEPEAPAPTTPEEPEAPAPTTPEEPEAPAAPATTPEEPEAPATTPEEPEAPAAPATTPEEPEAPAAPATTPEEPEAPAAPATTPEEPEAPAAPATTPEEPEAPAAPATTPEEPEAPAAPATTPEEPEAPAAPAAAAAAAA	S B	
73			
VT	3// TKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKE	Db V9	
-14	APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGV		
μį i	330 PKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTT	Qy	
- i	403 TKE-TARTITEKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTT	B 4	
- F	23DNKPAPGSTAPXAHG	Di Di	_
PT	Þ	Qy	
2 — 2 2 — 3 7 — 4		DЬ	
2	TKGPA	Qy	
S K	140 TSLTVNKETTVETKETTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT	Db Qy	
	Query Match 6.2%; Score 424.5; DB 4; Length 1867; Best Local Similarity 23.7%; Pred. No. 4.8e-20; Matches 255; Conservative 69; Mismatches 511; Indels 241;	Ques Best Mato	

```
; OTHER INFORMATION: ; OTHER INFORMATION: US-08-479-537A-5
                       δÃ
                                                           В
                                                                                            γO
                                                                                                                                 ф
                                                                                                                                                                Qy
                                                                                                                                                                                                          Вþ
                                                                                                                                                                                                                                         ρ
                                                                                                                                                                                                                                                                                    Matches 255;
                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: (703) 836-2021
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1867 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURE:
NAME/KEY: Peptide
LOCATION: 128..1727
OTHER INFORMATION: /r
OTHER INFORMATION: 12
OTHER INFORMATION: 12
OTHER INFORMATION: re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION UNMBER: 35,030
REFERENCE/POPULATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US 0: FILING DATE: 04-APR-1993 PRIOR APPLICATION DATA: APPLICATION NUMBER: US 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REGISTRATION NUMBER: 35,030
REGISTRATION NUMBER: 01
REFERENCE/DOCKET NUMBER: 01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: WO PCT/FR91/00835 FILING DATE: 23-OCT-1991
                                                                                                                                                                                                                                  140 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  OTHER INFORMATION:
OTHER INFORMATION:
                            283
                                                                                                                                                                      199 P-KAETTTKGP--ALTTPKEP------TPTTPKEPASTTPKEPTPTTIKSAP 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY: Peptide LOCATION: 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME/KEY: Peptide LOCATION: 1..21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         OTHER INFORMATION:
                                                                                                                                                                                                              16 TVLTV-----
       TKE--PAPTTT-----KSAPTTPKEPAPTTPK----KPAPTTPKEPAPTTPKEPTPTT 329
                                                             --DNKPAPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 180
                                                                                                    TTPKEPAPTTT-----KSAPTTPKEP--
                                                                                                                                   PGSGSSTTQGQDVTLAPATEPASGSAATWGQDVTSVPVTRPALGSTTPPAHDVT---SAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Peptide
134
                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                              --- VTGSGHASSTPGGEKETSATORSSVPSSTEKNAVSMTSSVLSSHS
                                                                                                                                                                                                                                                                                                         6.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note= "The amino acids spanning 128 to 1727 constitute a repeated region wherein the repeat 20 amino acids, 17 of which are fixed. The number of such repeats varies from 1 to 40."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US 08/039,320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Amino acid 147 is X2 = Xaa
which is the codon for Pro or Ala wherein Pro =
or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note= "Amino acid 144 is Y = Xaa which is the codon for Thr or Asn wherein Thr = or ACG; and Asn = AAT or AAC."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Amino acid 134 is X1 = Xaa which is the codon for Pro or Ala wherein Pro = CCT, CCC, or CCG; and Ala = GCT, GCC, GCA, or GCG."
                                                                                                                                                                                                                                                                                                                                                                                                /note= "Amino acids 1 to 21 are a 21 amino acid precursor sequence."
\underline{\underline{z}}^{\frac{1}{2}}
                                                                                                                                                                                                                                                                                          69;
                                                                                                                                                                                                                                                                                    Score 424.5; DB 2;
Pred. No. 4.8e-20;
9; Mismatches 511;
                                                                                                        APTITKEPAPTIPKEPAPTI 282
                                                                                                                                                                                                                                                                                                                             Length 1867;
                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                              241;
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                        65
                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CCT, CCC,
                                                                                                                                                                                                                                                                                                 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACI
```

Db	Qy	рь	ΟУ	В	ОУ	Дb	Qy	Db	Qy	рь	Qy	DЬ	Qy	Db	Qy	Db	Qy	Db	Qy	ДD	QУ	Д	Qy	DЪ	Qy	Db	Qy	Db
958	1008	912	949	869	889	821	829	763	773	703	727	645	683	586	637	531	7 589	472	534	420	478	360	426	301	377	241	330	181
TSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPD 1003	BETNICNGKPVDGLTT-LRNGTLVAFRGHYFWMLSPFSPPSPARRITEVWGIPSPID 1062	APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPX	-PNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMPEVTPDMDYLPRVPNQGIIINPMLSD	-GSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSA-PDXRPXPGST	KPTKAPKKPTSTKKPKTMPRVRKPKTTPTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQT	APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXP-	TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTEIMNKPEETAKPKDRATNSKATTPKPQ	DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS	-TKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS	DXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP 762		RPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAP 702	TPPPTTSEVSTP	PXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDX 644		TAPXAHGVTSAPDXRPXPGSTAP-XAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXR 585	PAPTTP	APXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRP-XPGSTAPXAHGVTSAPDXRPXPGS 530			PEEPAPT 	SAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 419		APDXRP-XPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVT 359		APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 300		APDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTSAPDXRPXPGSTAPXAHGVTS 240

RESULT 12
US-09-083-116-5
; Sequence-5, Application US/09083116
; Sequence-5, Application US/09083116
; Patent No. 6203795
; Patent No.

œ

```
APPLICANT: VAN HEERDE, GEORGE V.
APPLICANT: VAN HEERDE, ALEXIS C.
APPLICANT: BOUWSTEA, JAN B.
APPLICANT: BOUWSTEA, JAN B.
APPLICANT: DE WOLF, FREDERIK A.
APPLICANT: WERTEN, MARC W.T.
APPLICANT: WIND, RICHELE D.
APPLICANT: WIND, RICHELE D.
APPLICANT: VAN DEN BOSCH, TANJA J.
APPLICANT: VAN DEN BOSCH, TANJA J.
TITLE OF INVENTION: SUITWER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITWER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITWER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITWER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITWER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
TITLE OF INVENTION: SUITWER FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
TITLE OF INVENTION: WIJMBER: US/09/219,849
CURRENT APPLICATION NUMBER: US/09/219,849
CURRENT FILLING DATE: 1998-12-23
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PATENTAL OF SEQ ID NOS: 50
SOFTWARE: PATENTAL OF SEQ ID NOS: 50
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
                  밁
                                                       Ş
                                                                                                              밁
                                                                                                                                                      Ş
                                                                                                                                                                                                                                                                                                                                       ; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: Illustrative : OTHER INFORMATION: amino acid sequence US-09-219-849-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
US-09-219-849-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                   Query Match 6.6%; Score 452; DB 4; 1 Best Local Similarity 25.6%; Pred. No. 3.4e-22; Matches 173; Conservative 60; Mismatches 309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 5, Application US/09219849 Patent No. 6150081
127 GPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPAGPPGSRDPG 186
                                            255 APTTPKEPAPTTTKEPA-PTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTP---KKPA 310
                                                                                                                                              197 PTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAP--TTTKS 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            621 ҮРРТНКАКРТҮКАКРЅҮРРТНКАКР 645
                                                                                              72 PGPAGPPGSRDPG--PPGAPGPAGP--PGSRDPGPPGAPG-PAGPPGSRDPGPPGAPGPA 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         777 MTTTAKOKTTERDLRTTPETTTAAP 801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  664 APKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPA--PTTPKKPA-PTT----PETP 716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      606 EPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKGTAPTTLKEPA--PTTPKKP 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       553 EPA-PTT----PKEPAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAPKELAPTTTK 605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          363 AKPTYPSTYKAKPSYP----PTYKAKPSYPPTYKAKPSYPPTYK-AKPSYP----PTYKA 413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           309 PSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKYKPT--YK 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           442 PA--PTTPKEPA--PTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEK 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYK-AKPTYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKAKPTYPSTYKAKPSYP 576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTYKAKPSYPPT----YKAKPSYPPTYKAKPTYKA-----KPTYPST---YKAKPS 620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPE 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KPTYPSTYKAKPTYK--AKPSYP----PTYKAKPS-----YPPTYKAKPSYPPTYKAKP 521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KPTYPSTYKAKPSYPPTYKPKISYPPTYKAKPSYPSTYKAKSSYPPTYKAKPSYPPTYKA 473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LAPTTPE--KPAPTTPEELAPTTPEEPT--PTTPEEPA-PTTPKAAAPNTPKEPAPTTPK 552
                                                                                                                                                                                                                                                  Length 960;
                                                                                                                                                                                                   Indels 134;
                                                                                                                                                                                                Gaps
                                                                                                                                                                                                   41;
```

CLASSIFICATION: 514 *RIOR APPLICATION DATA: APPLICATION NUMBER: FR 90/13101 FILING DATE: 23-CCT-1990 *RIOR APPLICATION DATA:	יייייייייי
CATION NUMBE G DATE: 07-	
SYSTEM: Patent	
PUTER READABLE FO EDIUM TYPE: Flop OMPUTER: TAM PC	
STATE: Virginia COUNTRY: United States ZIP: 23313-1404	
P.O. Box 1404 Nexandria	••••
PONDENCE ADDRE	
NUMBER OF SEQUENCES: 5	
CANT: LATHE, Richard CANT: HAREUVENI, Mara	
•• ••	
11 9-537A-5 9-537A-5 0ce 5, App No. 5861 No. 5861	RESULT US-08-47 ; Sequen ; Patent ; GENER
696 SRDPGPPGAHGPAGPK 711	Ф
759 PKEPGVPTTKTPAATK 774	Qγ
PAGPPG 69	ф
ALENS 75	Qy
:: DPGPPGAPGPAGPPGSRDPGPPGA 6	ДЪ
~	Qy
:: PPGSRDP 58	DЬ
	Qy
63 GAHGPAGPKGAHGPAGPKGAHGPAGPK	DЬ
58 TPKEPAPTTPKET-APTTPKGTAPTTPKGTAPTTIKEPAPTTWEEDAPTT	Qy
OS - SRDPGPPGAPGPAGPPGSRDPGPGAPGPAGPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGS	₽ 5
AGPPGSRDPGPPGAPGPAGPPGSRDPGPAGPPG	Ou Db
ū	Qy
 PGAPGP 34	Ф
KKPAPTTPKEPAP 45	Qy
FAGPKGAHGPAGPKGAPGPAGPPGS	Db
A-DTTTKEDCDTTDKEDADTTTC 20	Q Y
7 PPGAPGPAGPPGSRDPGPPGAPGPAGPEGSRDPGPPGAHGPAGPKGAHGDAGPKGAHG	Db
311 PTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPA 349	Qγ

```
Ωy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 8
PCT-US92-00018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; TOPOLOGY: 1i
; MOLECULE TYPE:
PCT-US92-00018-2
                                                                                                                      В
                                                                                                                                                          Qy
                                                                                                                                                                                               뫄
                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                              망
                                                                        20
                                                                                                                                                                                                                                                                                                                QΥ
δÃ
                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: Spevack, AVTAM D.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEPAX: (301) 295-6759
TELEPAX: (301) 295-4033
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application PC/TUS9200018 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                       Query Match 7.2%; Score 488.5; DB 5; Best Local Similarity 26.4%; Pred. No. 1.1e-24; Matches 153; Conservative 54; Mismatches 223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-POS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: A. David
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Hedstrom, Richard
APPLICANT: Hedstrom, Richard
APPLICANT: Khusmith, Srisin
APPLICANT: Rogers IV, William O.
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria sporozoite surface protein
TITLE OF INVENTION: immunogen and gene encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     707
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        750 PTPKALENSPK--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            650 LPIIPQKGNNIPSNLPENPSDSEVEYPRPNDNGENSNNTMKSKKNT---PNEPIPSPGDN 706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               696 TPKEP-----APTTPKKPAPTTPETPPPTTSEVSTPTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                233 TPTTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTTKE 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STREET:
458 PAPTTTKKPAPTAPKEP-APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAP 516
                                       444 PNKPNP----NEPSNPNKPNPN-----EPLNPNEPSNPNEPSNPNAPSNPNE---PSNPNE 492
                                                                    400 TKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA-PTTPKKP-APTTPKEPAPTTPKE 457
                                                                                                                                             342 PTTPKEPAPTAPK--KDAPTTPKEPAPTTTKEPAPTTTKEPSPTTPKEPAPTTTKSAPTT 399
                                                                                                                                                                                                      333
                                                                                                                                                                                                                                       286 PAPTTTKSAPTTPKEP-APTTPKKP-APTTPKEP-APTTPKEP-TPTTPKEPAPTTKEPA 341
                                                                                                                                                                                                                                                                                 273 TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: FILING DATE: 19920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
ZIP: 20814-5044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PYKGHEERIPKPHRSNDDYVYDNNVNKNNKDEPEIPNNE 745
                                                                                                                      RRNPKRRNPNKPKPNKENPNKPNEPSNPNKPNPN-----EPSNPNKPNP----NEPSN 443
                                                                                                                                                                                                   Bethesda
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1: 826 amino acids
AMINO ACID
OGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Medical Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            E: A. David Spevack
NMRDC Building 1 T-12 National Naval
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Charoenvit, Yupin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hoffman, Stephen L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19920103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PCT/US92/00018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -EPGVPTTK 768
                                                                                                                                                                                                                                                                                                                                                                                                     Length 826;
                                                                                                                                                                                                                                                                                                                                                                    Indels 149;
                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            749
                                                                                                                                                                                                                                                                                                                                                                        34;
```

	395 SAPTITKEPAPTITKSAPTIPKEPSPTITKEPAPTIPK	
	Qy 344 TPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTKEPSPTTPKEPAPTTTK 394 	
	QY 289 TITKSAPITPKEPAPITPKKPAPITPKEPIPTITPKEPAPITKEPAPT 343 Db 139 PIYKPKPSYPPSYKTKKTYPPTYKPKLTYPPTYKPKPSYPPSYKPKKTYPPTYKPKLT 196	
	QY 232 PTPTTIKSAPT-TPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTTKEPAP 288	
	9 3	
	0	п 5
45;	Query Match 7.0%; Score 476.5; DB 6; Length 652; Best Local Similarity 29.7%; Pred. No. 5.1e-24; Matches 221; Conservative 71; Mismatches 284; Indels 169; Gaps	
	SEQ 1	رب _ا د. د.
	FILING DATE: 24-NO APPLICATION NUMBER FILING DATE: 13-SE	
	APPLICATION NUMBER: FILING DATE: 07-AUG-APPLICATION NUMBER:	
	NUMBER OF SEQUENCES: CURRENT APPLICATION I APPLICATION NUMBER FILING DATE: 25-MAY	
	RESULT 9 5202236-13 5202236-13 FRATENT NO. 5202236 APPLICANT: MAUGH, KATHY J.;ANDERSON, DAVID M.;STRAUSBERG, SUSAN L.;MCCANDLISS, RUSS;WEI, TENA;FILPULA, DAVID TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE PROTEIN	
	D 707 PYKGHEERIPKPHRSNDDYVYDNNVNKNNKDEPEIPNNE 745	Дb
	9y 750 PTPKALENSPK	Qy
	y 696 TPKEPAPTTPKKPAPTTPETPPTTTSEVSTPTTTKEPTT1HKSEDESTPELSAE /49	Db Qy
	636 PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT 69	Qy Db
	576 KGTAPITLKEPAPITPKKPAPKELAPITIKEPTSTTSDKPAPTTPKGTAPITPKEPAPIT 63	Qy Db
	517 TTPEEPTPTTPEEPAPTTPKAAAPUTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 5	Qy Db
	493 PSPSNPNEPSNPNEPSNPKK	DЬ

407

----TKEP----

PKE-----

Š В Š 밁 õ 밁 õ B õ 밁 õ

> 581 476 527 434 467

밁

681 596

Page

9

```
EVTPD-----MDYLPRVPNQGIIINP------MLSDETNICNGK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VYGLPVGEITGLPKDPGSDIPFNSTTGELVDPSTGKPINNSTAGIVSGKPGLPPIEDENG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IMKTPTQTDSVTGKPIDPTTGLPFNPPTGHLINPTNNNTMDSSFAGAYKYAVSNGIKTDN 1023
                                                                                                                                                                                                                                             QQTDEITGKPYDTYTGLPYDPSTGEIIDPATKLPIPGSVAGDEILTEVLNITTDEVTGLP
                                                                                                                                                                                                                                                                                                                   KDGLIVPPTNSINKDPVTNTQYSNTTGNI----INP--ETGKVIPGSLPGSLNYPSFNTP
                                                                                                                                                                                                                                                                                                                                                                                              NLFDPSTNLPIDGNNQLVNPETNSTVSGSTSGTTKPKPGIPVNGGGVVPDEEAKDQADKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IMNKP--EETAKPKD-----RATNSKATTPKPQKPTKAPKKPTSTKKPKTMP-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TKETATTTEKTTESKITATT-TQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPT----TTKEPTTIHKS-----
                                                                                                                                                                                                                                                                                                                                                        -----LNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TNMVVGEFGGLLNPATGVMIPGFLGPSEQTQFSPEIEDGGIIPPEVAAANADKFKLSIP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LTGYPLDPVSLIPFN--PETGELFDPISDEIMNGTIAGIVSGISASESLLSQKSALIDPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KSTGLPIDPMVGLPFDPKSGNLVHPYTNQTMSGLSVSYLAAKNLTVDTDETYG--LPIDT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -PSVP-----ESIPE-----KDQKIDSISELMYDIESGRLIGQVSKRPIPGSIAGDLNP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKEPAPTTPKGTAPTTLKE---PAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --KVPGKPPIATTTTLKPIVTTTTTKATTTTTTTVPTT----TTTTKRDEMTTTTTPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KAAAPNTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPA 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----PELSAEPTPKALENSPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDIGDIEITPIPIEKMLDKYTRMIYDYNSGLLLDSNDEPIPGSQAGQIADTSNLFPVQTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PK----ELAPTTTKEPTS--------TTSDKPAPTTPKG-TAPTTPKEPAPTT 635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PGQPTTTTTGSPSKPTTTTTTKATTTT-----TTLNPIITTTTQKPTTTT-----TT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VSLIS----APIQPSELFNEVYCDTCTAKYGAIHSGYQTSADFVTTTTAKPTTTT--TGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VKDVGNTISVRCRKGAGKLEFPDRSLDFTIPPVAGHNSCSIIVGVSGDGKIHVSPYGSKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSETESVIKPDEWCWLEKNGECEAKGATYVGVIGKDGRIENGMAFTMIPNDDTHVRFRFK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----DLRTTPETTTAAPKM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----PSPTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----APTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1083
                                                                                                                                                                                                                                                                                   1015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     692
                                                                                                                                                                                                                                                                                                                                                                                                                                       927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    422
                                              δõ
            В
                                                                                                                       Qy
                                                                                                                                                              B
                                                                                                                                                                                              Š
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TYPE: protein US-07-638-431-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: (301) 295-403
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Spevack, Avrom D. TELECOMMUNICATION INFORMATION: TELEPHONE: (301) 295-6759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compati
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: A. David Spevack
STREET: NMRDC Building 1 T-12 National
STREET: Medical Center
CITY: Bethesda
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Hedstrom, Richard
APPLICANT: Khusmith, Srisin
APPLICANT: Rogers IV, William O.
TITLE OF INVENTION: Protective malaria sporozoite surface
TITLE OF INVENTION: immunogen and gene
          607
                                              636
                                                                                                                                                                524
                                                                                                                                                                                                    517
                                                                                                                                                                                                                                        493
                                                                                                                                                                                                                                                                           458
                                                                                                                                                                                                                                                                                                                 444
                                                                                                                                                                                                                                                                                                                                                        400
                                                                                                                                                                                                                                                                                                                                                                                            393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         233 TPTTIKSAPTTPKEPA-PTTTKSAPTTPKEPA-PTTTKEP----APTTPKEP-APTTTKE 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: FILING DATE: 19910 CLASSIFICATION: 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                      KGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTT 635
                                  PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPT 695
                                                                                                                                                 SNPNE--PLNPNEP-----SNPNEPSNPNEPSNPEE--PSNPKE--PSNPNE-----
                                                                                                                                                                                    TTPEEPTPTTPEEPAPTTPKAAAPNTPKEP-APTTPKEPAPTTPKEPAPTTPKETAPTTP 575
                                                                                                                                                                                                                                                                                                     PNKPNP---NEPSNPNKPNPN-----EPLNPNEPSNPNEPSNPNAPSNPNE--PSNPNE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAPTTTKSAPTTPKEP-APTTPKKP-APTTPKEP-APTTPKEP-TPTTPKEPAPTTKEPA 341
          PKEPI---
                                                                                                                                                                                                                                                                 PAPTTTKKPAPTAPKEP-APTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAP 516
                                                                                                                                                                                                                                                                                                                                          TKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA-PTTPKKP-APTTPKEPAPTTPKE 457
                                                                                                                                                                                                                                                                                                                                                                              RRNPKRRNPNKPKPNKPNPNEPSNPNKPNPN----EPSNPNKPNP----NEPSN 443
                                                                                                                                                                                                                                                                                                                                                                                                                   PTTPKEPAPTAPK--KPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                           TPCKVRDCPQIPIPPVIPNKIPEKPSNPEEPVNPNDPNDPNNPNNPNNPNNPNNPNNPNNPNN 332
                                                                                                                                                                                                                                  PS-----NPNEPSNPNEPSNPNE--PSNPKK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4: 826 amino acids
AMINO ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20814-5044
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7.2%; Score 488.5; llarity 26.4%; Pred. No. 1.1e Conservative 54; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19910110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1) 220
295-4033
--- NO: 2:
-NPEESNPKEP----INPEDNENPLIIQDEPTEPRNDSNVIPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US/07/638,431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5; DB 1;
1.1e-24;
nes 223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Naval
                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
649
                                                                                                                                                   564
                                                                                                                                                                                                                              523
                                                                                                                                                                                                                                                                                                         492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34;
```

RESULT 7
US-07-638-431-2
Sequence 2, Application US/07638431
Patent No. 5198535
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charcenvit, Yupin

밁 Š 밁 Š

IDLETGLPRDPVSGLPQLPNGTLV 1281 -----PVDGLTTLRNGTLV 1029 1198

Š 8

928

1144

밁 õ 밁 Š В Ş В Š 밁 S 밁 S 밁 Š

1084

1024

863 964 804 761 857

EPGVPTTKTPAATKPEMTTTAKDKTTER------

799

741

908

----RVRKPKTTPTPRKMTS--TMPE----

```
밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DЬ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             р
                                                                                                                                                                                                                                 망
                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                    ρy
                                                                                                                                                                                                                                                                                                                                                   Дb
                                                                                                                                                                                                                                                                                                                                                                           γQ
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DЬ
                                                                                                                                           Qy
                                                                                                                                                                        рь
                                                                                                                                                                                                  QΥ
                                                                                                                                                                                                                                                             δÃ
В
                                                       DЬ
                                                                                 δō
                                                                                                                 Вb
                                                                                                                                                                                                                                    1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     467 VKDVGNTISVRCRKGAGKLEFPDRSLDFTIPPVAGHNSCSIIVGVSGDGKIHVSPYGSKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           415
                                                                                                                                                                                                                                                                                                                                                      916
                                                                                                                                                                                                                                                                                                                                                                                                             857
                                                                                                                                                                                                                                                                                                                                                                                                                                          739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   693
                                                                                                                                                                                                                                                              863 IMNKP--EETAKPKD-----RATNSKATTPKPQKPTKAPKKPTSTKKPKTMP-----
                                                                                                                                                                                                                                                                                             964
                                                                                                                                                                                                                                                                                                                                                                                761
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --KVPGKPPIATTTTLKPIVTTTTKATTTTTTVPTT----TTTTKRDEMTTTTTPL 680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPA 595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PGOPTTTTTGSPSKPTTTTTTKATTTT-----TTLNPIITTTTTTTQKPTTTT-----TT- 627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTP 535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VSLIS----APIQPSELFNEVYCDTCTAKYGAIHSGYQTSADFVTTTTAKPTTTT--TGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPTTPKKPAPTTPKE-------PAPTTPKEPAPTTTKKPAPTAPKEPA 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSETESVIKPDEWCWLEKNGECEAKGATYVGVIGKDGRIENGMAFTMIPNDDTHVRFRFK 466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LTGYPLDPVSLIPFN--PETGELFDPISDEIMNGTIAGIVSGISASESLLSQKSALIDPA 856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KSTGLPIDPMVGLPFDPKSGNLVHPYTNQTMSGLSVSYLAAKNLTVDTDETYG--LPIDT 798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDIGDIEITPIPIEKMLDKYTRMIYDYNSGLLLDSNDEPIPGSQAGQIADTSNLFPVQTH 740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PK----ELAPTTTKEPTS--------TTSDKPAPTTPKG-TAPTTPKEPAPTT 635
                                                                                                                                                                                                                                                                                           IMKTPTQTDSVTGKPIDPTTGLPFNPPTGHLINPTNNNTMDSSFAGAYKYAVSNGIKTDN 1023
                                                                                                                                                                                                                                                                                                                       TKETATTTEKTTESKITATT-TQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE 862
                                                                                                                                                                                                                                                                                                                                                                                {\tt TNMVVGEFGGLLNPATGVMIPGFLGPSEQTQFSPEIEDGGIIPPEVAAANADKFKLSIP-\\
                                                                                                                                                                                                                                                                                                                                                                                                                                     ------PELSAEPTPKALENSPK 760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APTTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPT-----TTKEPTTIHKS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKEPAPTTPKGTAPTTLKE---PAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKET 692
                                                        QQTDEITGKPVDTVTGLPYDPSTGEIIDPATKLPIPGSVAGDEILTEVLNITTDEVTGLP 1257
                                                                                                                 KDGLIVPPTNSINKDPVTNTQYSNTTGNI----INP--ETGKVIPGSLPGSLNYPSFNTP 1197
                                                                                                                                                                          NLFDPSTNLPIDGNNQLVNPETNSTVSGSTSGTTKPKPGIPVNGGGVVPDEEAKDQADKG 1143
                                                                                                                                                                                                                                   VYGLPVGEITGLPKDPGSDIPFNSTTGELVDPSTGKPINNSTAGIVSGKPGLPPIEDENG 1083
                                                                                                                                                                                                                                                                                                                                                    -PSVP-----ESIPE-----KDQKIDSISELMYDIESGRLIGQVSKRPIPGSIAGDLNP
                                                                                  EVTPD-----MDYLPRVPNQGIIINP-------MLSDETNICNGK-----
                                                                                                                                             ----LNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRPHVFMP
                                                                                                                                                                                                     -----RVRKPKTTPTPRKMTS--TMPE----
 IDLETGLPRDPVSGLPQLPNGTLV 1281
                             ------PVDGLTTLRNGTLV 1029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------PSPTT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       350
                                                                                                                                                      982
                                                                                          1015
                                                                                                                                                                                                                                                                       907
                                                                                                                                                                                                                                                                                                                                                                                                                   915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         738
```

RESULT

σ

406		B 2	
414	CT CONTENTS TO THE TOTAL T	, ,	
350	230 TTTTKKPTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	p 5	
354	Adella Valla value de la Valla V	? !	
300	235 TTIKSAPTIPKEPAPTTTKSAPTIPKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTKEPAPTTTTTTTTTT	Оy	
240		Дb	
234	TPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTP	Qy	
208	ENSFEQGQIFDMGSKVYIPYTKCVGVKHTTTTTTTTTTTTTTTTTT	Db	
174	GKEKTTSAK	Qy	
158		Дb	
129	TAKPINPRPS	Qy	
ps 41;	Query Match 7.2%; Score 489; DB 3; Length 1721; Best Local Similarity 22.1%; Pred. No. 2.4e-24; Matches 284; Conservative 82; Mismatches 476; Indels 442; Gaps	Ma Be	
	MOLECULE TYPE: protein 08-928-361B-6	, , ,	
	STRANDEDNESS: TOPOLOGY: linear		
	LENGTH: 1721 amino acids		
	INFORMATION FOR SEQ ID NO: 6: SEQUENCE CHARACTERISTICS:		
	TELEPHONE: 650-324-1677 TELEFAX: 650-324-1678		
	REFERENCE/DOCKET NUMBER: 400./0-1(AV) TELECOMMUNICATION INFORMATION:	٠. ٠.	
	518	., .,	
	ATTORNEY/AGENT INFORMATION:	٠. ٠	
	APPLICATION NUMBER: US 60/026,062 FILING DATE: 13-SEP-1996		
		٠, ٠.	
	FILING DATE: 12-SEP-1997		
	A:	•. •	
	OPERATING SYSTEM: PC-DOS/MS-DOS COPTWARE: Patentin Release #1.0, Version #1.30		
	TYPE: Floppy disk R: IBM PC compatible		
	ADABI		
	COUNTRY: USA	·· ··	
	CITY: Palo Alto	٠.	
	ADDRESSEE: PETERS, VERNY, JONES & BIKSA STREET: 385 Sherman Avenue, Suite 6	 .	
	NUMBER OF SEQUENCES: 30 CORRESPONDENCE ADDRESS:		
	TES INFECTIONS		
GS AND FRAGME	TITLE OF INVENTION: PEPTILDES, FOLIEBET HUTANTS, ANALOGS AND TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, ANALOGS AND THE OF INVENTION OF THE PROPERTY OF T		
	APPLICANT: Petersen, Carolyn		
	Patient No. 6 17:F118 Patient No. 6 17:F118 Patient Transportation.	; Pa	
	US-08-928-361B-6 : Sequence 6, Application US/08928361B	. Sec.	

```
밁
                                                  õ
                                                                                              В
                                                                                                                                  Ş
                                                                                                                                                                              밁
                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                               밁
                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                 DЪ
                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HYPOTHETICAL: NO
ORIGINAL SOURCE: NO
ORIGINAL STATE: Peritrophic membrane
US-09-103-429A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE/DOCKET NUMBER: BTI-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (607) 256-2000
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 786 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET: 118 NO. 6187558th Tioga
CITY: Ithaca
STATE: NY
COUNTRY: USA
ZIP: 14850
COMPUTER READABLE FORM:
MEDLIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PAtentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION UNMBER: US/09/103,429A
EILING DATE: 24-JUN-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Chriscopher A
REGISTRATION NUMBER: 34,390
DEFERENCE/DOCKET NUMBER: BTI-39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match 7.3%; Score 498.5; DB 4; Length 786; Best Local Similarity 29.1%; Pred. No. 2.2e-25; Matches 200; Conservative 25; Mismatches 207; Indels 255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 3, Application US/09103429A Patent No. 6187558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GENERAL INFORMATION:
362 CGAGTHFSFELOOCDHIELVGCTLPGGESEEVDVDEDA-CTGWYCPTEPIEWEPLPNGCP 420
                                                                             302 DGEISPAPPVTEGNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERR 361
                                                                                                                                                              242 ELLPNGCPADFDIHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVEC 301
                                                                                                                                                                                                                          374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin TITLE OF INVENTION: cDNA and Related Products and Methods NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS: ADDRESS: Brown, Pinnisi & Michaels, P.C.

STREET: 118 No. 6187558th Tioga
                                                                                                                                    387 ----EPAPTTT-----
                                                                                                                                                                                                                                              184 -APTITQAPTTITQAATTPAATTPAATTPAATTPAA-TTPAATTPGVPAPTSAPVWPPIC 241
                                                                                                                                                                                                                                                                           317 PAPTTPKEPTTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTT-PKEPA-- 373
                                                                                                                                                                                                                                                                                                                             136 TTTQ--APTTTQ--ATTT---QAPTTTTQ-APTTTTQAPTTTTQ-APTTTQ--APTTTQ- 183
                                                                                                                                                                                                                                                                                                                                                                                                                                       197 PTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAP 256
                                                                                                                                                                                                                                                                                                                                                      257 TTPKEPAPTTTKEPAPTTPKEPAPTTTKSAPTTPKEPAPTTPKKEPAPTTPKKE 316
                                                                                                                                                                                                                                                                                                                                                                                                               93 PGPPAE-TYQAPATT----QAPTTTQAPTTTT------QAPTTTTQ--ATTTTQAP 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Granados, Robert R
APPLICANT: Wang, Ping
                                                                                                                                                                                                      386
                                     -----KSAPTTIKEPAPTTTKSAPTTPKEPSPTTTKEP 426
                                                                                                                         .---- 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32;
```

	PTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPAPTTPKEPAPTTPKEPAPAPTTPKEPAPTTPKEPAPAPTTPKEPAPTTPKEPAPTTPKEPAPAPAPTTPKEPAPAPTTPKEPAPAPTTPKEPAPAPAPTTPKEPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP	Qy 295	
	▗▗▗▗▗▗▗▗ ▃▃▗▃▗) 1	
	INDEPENDENT TENSOR IN THE TENS		
	O 77 LILLITITITITITITITITITITITITITITITITITI	0v 235	
		Db 209	
	ETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTP 2	Qy 175	
	APCNSENSFEQGQIFDMGSKVYIPYTKCVGVKHTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Db 159	
	LPPNSDTSKETSLTVNKETTVETKETTTINKQTSTDGKEKTTSAK	ΩУ 130	
		Db 116	
	0 KKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKITTAKPINPRPS 129	Ωу 7	
	/ Match 7.2%; Score 489; DB 3; Length 1721; Local Similarity 22.1%; Pred. No. 2.4e-24; Les 284; Conservative 82; Mismatches 476; Indels 442; Gaps 41;	Query M Best Lo Matches	
	08-700-651-5	-S	
	75		
	NRE: PatentIn Ver. 2.0	SEO	
	REFILING DATE: 1995-04-03	EAR	
	APPLICATION NUMBER: US/08/7 FILING DATE: 1997-08-14 APPLICATION NUMBER: 08/115	CURRENT	
•	INVENTION: INFECTIONS ERENCE: 480.19-4(HV)	FII	
RNAS	INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNA		
	APPLICANT: LEECH, JAMES APPLICANT: LETT TITE APPLICANT: CITT TITE		
	ENERAL INFORMATION: APPLICANT: PETERSEN, CAROLYN		
	0.5	i	
	700-651-5	RESULT	
	78APATTPEDDDIDPPLPNDPIND 699	Db 6	
		0у 6	
	ഗ —	Db 6	
	TPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTS	Оу б	
		Db 5	
	SDKPAPTTPK	Qy 5	
	: VPEIPITVTSAPTAAPT	Db 5	
	PEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKE	Qу 5	
		Db 4	
	י ס	Qy ,	
	GCOVEECDSD	, שמ	
	42/ APTTPKEP 446	ζ.	

RESU 5202 5202 ; Pat ; SUS ; PRO	Фу	ОУ	pb Qy	ОУ	Фу	Оу	ОУ	Оу	Qу Db	Оу	ОУ	Qy Db	Оу	Оy	Оу
RESULT 3 202236-25 202236-25 Patent No. 5202236 Patent No. 5202236 PATELICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG, SUSAN L.; MCCANDLISS, RUSS; WEI, TENA; FILDULA, DAVID TITLE OF INVENTION: METHOD OF PRODUCING BIOADHESIVE PROTEIN NUMBER OF SEQUENCES: 39 NUMBER OF SEQUENCES: 39 APPLICATION NUMBER: US/07/528,762 FILING DATE: 25-MAY-1990 APPLICATION NUMBER: 82,456	1032 RGHYF 1036 1637 SGKYF 1641	1002 NPMLSDETNIC	964	9 WW 1	863IMNKPEETAKPKDRATNSKATTPKPQKP-TKAPKKPTSTKKPKTMPRVRKPKT 914	804 TKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTE- 862	М 80 : 1 13	1	667ELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPA 701	624 APTTKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPK 666		10	A 53	491 T 504	477KKL 490

```
FILING DATE: 07-AUG-1987
APPLICATION NUMBER: 933,945
FILING DATE: 24-NOV-1986
APPLICATION NUMBER: 650,128
FILING DATE: 13-SEP-1984
SEQ ID NO:25:
LENGTH: 744
5202236-25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΨ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       рb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                         QΥ
                                                                                                                                               Вþ
                                                                                                                                                                                 Qy
                                                                                                                                                                                                                        рь
                                                                                                                                                                                                                                                        Ωy
                                                                                                                                                                                                                                                                                               DЪ
                                                                                                                                                                                                                                                                                                                               Qγ
RESULT 4
US-09-103-429A-3
                                                                        Ъ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 7.4%; Score 506.5; DB 6; Length 744; Best Local Similarity 29.9%; Pred. No. 6.2e-26; Matches 239; Conservative 78; Mismatches 308; Indels 175;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            116 PKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKE 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       152 PTYKPKPSYP--PSYKTKKTYPPTYKPKLTYPPTYKPKPSYPPSYKPKKTYPPTYKPKLT 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               176 TQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKE----PASTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    341 -APTTPKEPA--PTAPKKPA--PTTPKEPA--PTTPKEPA--PTTTKEPS--PTTPKEPA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    232 PTPTTIKSAPT-TPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPKEPAPTTTKEPA--P 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           390 -PTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPA--PTTPKEPA----PTTPKKP 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              210 YPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPS 269
                                                                                                                                                                                                                                                                                                                                                                            499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          383 TYKAKPSYPPTYKAKPSYP----PTYKAKPSYPPTYKAKPSYPPTYKAKPTYPSTYKAKP 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           495 PEKLAPTTPE---KPAPTTPEELAPTTPEEPT---PTTPEEPA-PTTPKA--AAPNT-----P 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 323 SYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKVKPTYKAKPTYPS 382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 -----YPPTYK--AKPS-----YPPTYKPKTYPPTYKPKLTYPPTYKPKP 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24 PKMTYPPTYKPKPSYPP-----KPKIT----TYKSKPTY-----KPKIT----
                                                                                                                                                                                                                                                                                                   547
                                                                                                                                                                                                                                                                                                                       657 -PTTPKKPAPKELAPTTTKGPT--STTSDKPAPTTPKETAPTTPKEPAPTTPKKPA--PT 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  544 KEPAPTTPKEPAPTT----PKEPAPTTPKETAPTT--PKGTAPTTLKEPAPTTPKKPAPK 597
                                                                                                                                                                                                                        602 YKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAK 661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93
                                                                                                                                               662 PSYP--PTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAKPSYPPTHKAKPT 719
                                                                                                                                                                                                                                                          712 ---TPETPPPTTSEVSTPTTTKE----PTTIHKSPD-----ESTPELSAEPT-PKALENS 758
                                                                                                              805 KETATTTEKTTESKITATTT 824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTTKSAPTTPKEPAPTTPKKPAPT-TPKEPAPTT----PKEPTPTTPKEPAPTTKEP--- 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SYPPTYKSKPTYKPKITYPPTYKAKPSYPPTYKPKTYPPT-YKPKLTYPPTYKPKASYP 151
                                                                                                                                                                                                                                                                                                 PPTYKAKPTYK-AKPTYKAKPTYPSTYKAKPSYSPTYKAKPSYP-----PTYKAKPSYPPT
                                                                                                                                                                                                                                                                                                                                                                        SYPPTYKAKPTYPSTYKAKPTYK--ARPSYP----PTYKAKPS-----YPPTYKAKPSY 546
                                                                                                                                                                                                                                                                                                                                                                                                         ELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPA- 656
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYPPTYKPKISYPPTYKAKPSYPSTYKAKSSYPPTYKAKPSYPSTYKAKSSYPPTYKAKP 498
                                                                                                                                                                                     PKEPGVPTTKTPAA-----TKPEMTTTAKDKTTERDLRTTPETTTAAP------KMT 804
                                                                          YKAKPTYPSTYKAKPTYPST
                                                                              739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53
                                                                                                                                                                                                                                                                                                               601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         50;
```

N

```
RESULT 2
US-08-928-3618-5
                                                                                                                                                                  Sequence 5, Application US/08928361B Patent No. 6071518
                                                                                                                                                    GENERAL INFORMATION:
                                                                      TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                 CORRESPONDENCE ADDRESS:
                                                                                                                                       APPLICANT:
                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                         686 NTTAAP-----PTTTSAPATT-PE
                                                                                                                                                                                                                                                                                     745 ELSAEPTPKALENSPKEPGVPTTKTPAATKPE 776
                                                                                                                                                                                                                                                                                                                             642
                                                                                                                                                                                                                                                                                                                                                           583
                                                                                                                                                                                                                                                                                                                                                                                           599
                                                                                                                                                                                                                                                                                                                                                                                                                         626
                                                                                                                                                                                                                                                                                                                                                                                                                                                         549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    434 DCGQYLQCVHGQTIARPCPGNLHFSPATQSCESPVTAGCQVFECDSDNQCTSTAAPTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              314 GNEDEDIDIGDLLDNGCPANFEIDWLLPHGNRCDKYYQCVHGNLVERRCGAGTHFSFELQ 373
                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            254 IHLLIPHDKYCNLFYQCSNGYTFEQRCPEGLYFNPYVQRCDSPANVECDGEISPAPPVTE 313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          199 TTPAATTPAAT-TPAATTPAATTPAATTPGVPAPTS----APVWPPICELLPNGCPADFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      353 PKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPTTTKEPAPTTTKSAP 412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    293 SAPTTEKEPAPITEKKPAPTTEKEPAPTTEKEPTETTEKEPAPTTKEPAPTTEKEPAPTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67
                                                                                                                                                                                                                                                                                                                AP----NTTVTAPPTAAPTT-AAPAPNTTVTVPPTAAPTAAP-----PTVAH-----AP
                                                                                                                                                                                                                                                                                                                                                                                                TTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTST-TSDKP
                                                                                                                                                                                                                                                                                                                                       APTTEKETAPTTEKEPAPTTEKKPAPTTPETPEPTTSEVSTPTTTKEPTTIHKSPDESTP
                                                                                                                                                                                                                                                                                                                                                                                 TTAVPEIPTT-----VTSPPTAAPTT-AAPAPNT------
                                                                                                                                                                                                                                                                                                                                                                                                                                            PESPTTVTVPPTAAPTA----APTT----AVPEIPITVTSAPTAAPT--AAPTAAPTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ETAPTT----PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TAAPTAAPTAAPTAAPSTVVPPA-TPPATAAPVPPTTAIPT----PAPTAAPTAAPTTAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QCDHIELVGCTLPGGESEEVDVDEDACTGWYCPTEPIEWEPLPNGCPADFSIDHLLPHES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TTPKEPSPTTTKEPAPTTP--KEPAPTTPKKPAPTTPKEPAPTTPK------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---QAPTTTTQ-APTTTTQ-APTTTQ------APTTTQ-APTTTQ-APTTTQAA 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QAPTTTQATTTT-----QAPTTTQAPTTT----QATTTQAPTTTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RDCAPGTEFKFSAQTCVHAALAGCTLPGPPAETT---QAPATTQ--APTTTQ--APTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEPAPTT-----TKSAPTTPKEPAPTTTKEPAPTTTKEPAPTTTK 292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TALGLVAARPEVSDAEKNPALHEPHPDXP---PAEQXXLLPXEYDCTKFYYCEYGLKFIAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAP-----TTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226;
      385 Sherman Avenue,
                                                                                                                                  Petersen, Carolyn
     PETERS, VERNY, JONES & BIKSA
5 Sherman Avenue, Suite 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----PA------
                                                                 PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
FOR TREAFMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
SPECIES INFECTIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---APTTPKKLTPTT-----PEKLAPTTPEKPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----PTAP--KEPAPT----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                          703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----EPAPTTTK 464
                                                                                                                                                                                                                                                                                                                                                                               -TVTVPPTAAPTTAAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---TPKET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   342;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  482
                                                                                                                                                                                                                                                                                                                   685
                                                                                                                                                                                                                                                                                                                                                744
                                                                                                                                                                                                                                                                                                                                                                               641
                                                                                                                                                                                                                                                                                                                                                                                                              684
                                                                                                                                                                                                                                                                                                                                                                                                                                            598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       352
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40;
           B
                                    Qγ
                                                                         망
                                                                                                    Ş
                                                                                                                                     멍
                                                                                                                                                                  δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRANDEDNESS:
TOPOLOGY: lin
MOLECULE TYPE:
US-08-928-361B-5
                                                                                                                                                                                                   DЬ
                                                                                                                                                                                                                              ρy
                                                                                                                                                                                                                                                                 Вb
                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                            Дb
                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       В
                                                                                                                                                                                                                                                                                                                                                                                          멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: US
FILING DATE: 12-SEP-19
CLASSIFICATION:
PRIOR APPLICATION DATA:
                                         452
           782
                                                                         727
                                                                                                                                     667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana
                                                                                                                                                                     357
                                                                                                                                                                                                   607
                                                                                                                                                                                                                                 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEPHONE: 650-324-1677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CADRICATION DATA:
                                                                                                                                                                                                                                                               547
                                                                                                                                                                                                                                                                                              325
                                                                                                                                                                                                                                                                                                                            487
                                                                                                                                                                                                                                                                                                                                                          276
                                                                                                                                                                                                                                                                                                                                                                                          427
                                                                                                                                                                                                                                                                                                                                                                                                                                                      367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                        218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98 KYTTPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTT 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 21.3%; Pred. No. 1.4e-25; es 295; Conservative 100; Mismatches 494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 30 REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COUNTRY:
TTTTKRDEMTTTTTPLPDIGDIEITPIPIEKMLDKYTRMIYDYNSGLLLDSNDEPIPGSQ
                                                                                                                          AKYGAIHSGYQTSADFVTTTTAKPTTTTTGAPGQPTTTTTTGSPSKPTTTTTTKATTTTTI
                                      PTTPKEPAPTTTKKPAP----
                                                            LNPIITTTQKPTTTTTKVPGKPPIATTTTLKPIVTTTTKATTTTTTVPTT-----
                                                                                        ----APTITIKEPAPTITIKSAPTIPKEPSPTITKEPAPTIPKEPAPTIPKKPAPTIPKEPA 451
                                                                                                                                                                                      RSLDFTIPPVAGHNSCSIIVGVSGGGKIHVSPYGSKDVSLISAPIOPCELFNEVYCDTCT
                                                                                                                                                                                                                                                  AKGATYVGVIGKDGRIENGMAFTMIPNDDTHVRFRFKVKDVGNTISVRCGKGAGKLEFPD 606
                                                                                                                                                                                                                                                                                                              TTTATTTTKKPTTTTTTTTTTKKPTTTTTATTTTTSETESVIKPDEWCWLEKNGECE 546
                                                                                                                                                                                                                                                                                                                                            KEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKE------
                                                                                                                                                                                                                                                                                                                                                                            PTTPKEPASTTPKEPTPTTIKSAPTTPKEP---APTTTKSAPTTPKEPAPTTTKEPAPTTP 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPT 217
                                                                                                                                                                                                                                                                                                                                                                                                                                          94306-1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1837 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                    ----APTTPKEPAPTT---PKEPAPTTTKEPS-PTTPKEPAPTTTKS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12-SEP-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.5%; Score 508; DB 3; 21.3%; Pred. No. 1.4e-25;
                                                                                                                                                                                                                                                                                     US/08/928,361B
                                                                                                                                                                                                                      ------APTTPKEPA-PTAPKKP-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30,518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480.76-1(HV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Version
                                 TAPKEPAP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          #
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          . 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1837;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels 496;
                                                                                                                                                                                                                                                                                   ----PK 331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
 841
                                                            781
                                                                                                                                                        395
                                                                                                                                                                                      666
                                                                                                                                                                                                                      356
                                                                                                                                                                                                                                                                                                                                              324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      48;
```

₽ Š 멍 Š 8 Š 망 Š 망 Ş 당 Š В Ş 밁 Š 밁 ş 밁 Ş 밁 Š 밁 Ş 망 Ş 밁 Ş

```
Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OM protein -
                                                                                                                                                                                                                                                                                                                                                                   Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Minimum
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB seq
                                                                                                                                                                                                                                    489
488.5
488.5
476.5
                                                                                                                                                                                                                                                                                                                                                                    Score
                                                                                                                                                        424.5
424.5
419.5
                                                                                                                                                                                             424.5
                                                                                                                                                                                                         424.5
                                                                                                   419
417
417
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/
2: /cgn2_6/ptodata/2/
3: /cgn2_6/ptodata/2/
                                                                                                                                                                                                                                                                                                                                                                    Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 April 26, 2002, 16:25:33; Search time 61.21 Seconds (without alignments) 466.904 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         212252 seqs, 22503292 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAWKTLPIYLLLLSVFVIQ......ARAITTRSGQTLSKVWYNCP 1270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd
            /cgn2_6/ptcdata/2/iaa/5A_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/5B_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/6A_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/6B_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/PCTUS_COMB.pep:*
/cgn2_6/ptcdata/2/iaa/backfiles1.pep:*
                                                                                                                                                                                                                                                                                                                                                                    Length
786
1721
1721
826
826
826
652
960
1867
1867
1887
1883
837
837
837
837
837
837
                                                                                                                                                                                                                                                                                                                                                                        DB
            US-09-219-849-5
US-08-479-537A-5
US-08-479-537A-2
US-09-083-116-2
US-09-083-116-2
US-09-0841-886-3
US-08-642-255-132
US-08-642-255-132
US-08-477-509B-103
US-08-477-509B-103
US-08-477-509B-103
US-08-482-085B-103
US-08-482-085B-103
US-08-397-633A-50
US-08-397-633A-50
US-08-397-633A-50
US-08-397-633A-50
US-08-397-633A-50
US-08-783-774-2
US-08-783-774-2
US-08-783-774-2
                                                                                                                                                                                                                                                            US-09-103-429A-3

US-08-7.00-651-5

US-08-928-361B-6

US-07-638-431-2
                                                                                                                                                                                                                                                                                                                  US-09-103-429A-4
US-08-928-361B-5
5202236-25
                                                                                                                                                                                                                                                   PCT-US92-00018-2
    US-07-609-716-65
                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                      Sequence 3, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
                                                                                                                                                                                                                                                                                                                 Sequence 4, Appli
Sequence 5, Appli
Patent No. 5202236
                                                                                                                                                                                                                                                                                                                                                                        Description
    sequence 2, Appli
patent No. 5202236
Sequence 5, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 23, Appli
Sequence 23, Appli
Sequence 132, Appli
Sequence 132, Appli
Sequence 131, Appl
Sequence 101, Appl
Sequence 101, Appl
Sequence 103, Appl
Sequence 104, Appl
Sequence 105, Appl
Sequence 106, Appl
Sequence 107, Appl
Sequence 108, Appl
Sequence 108, Appl
Sequence 109, Appl
Sequence 109, Appl
Sequence 109, Appl
                                                                                                                                                                                     Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                      Patent No.
```

45	44	43	, i	3	41	40	L.	, (ا بد	37	36	Ü	ы Л	34	(L) (u S	ω L	30	ינ	29	ά)
367.5	367.5	367.5	,	368	368	370.5	3/1	2 1	371	371	374.5		275 5	376	307	201	381	398.5	230.7	3 0 0	404.5	404.0	n
5.4	5.4		1	5.4	5.4	5.4	0.4	n !	5 4	ъ	U	1 (л Л	ს		ית	თ	œ		n D	9	0.3	n 0
1187	7.81	1 - 0 -	1107	762	762	1064	100	763	762	761	LDS/	ון נו	1231	960	10	18/8	1848	789	100	CAN	408	4	108
U	·	+ د	_	\vdash	ب	-	٠,	٠	\vdash	Ν	H	د	w	4	. ر	л	4	+	۱ د	_	4		^
PCT-US94-U4496-28	02 - 040 - 040 - 20	110 00 0E0 0E0 = 28	115-08-320-559-28	US-08-397-633A-31	US-08-642-255-120	US-U8-046-650-66		11C-08-397-633A-26	US-08-642-255-114	US-08-/0/-23/A-04	CO-CO CAC ACCAS &	TE-09-305-367A-2	US-08-904-263A-4	US-04-219-849-6	TO 000 000 000 000 000 000 000 000 000 0	PCT-IIS95-10661A-6	US-08-296-791-6	US-U8-39/-033A 30	36 - 46231-36	115-08-642-255-126	US-U8-4/8-029A-02		IIS-08-475-411A-65
		28,	Sequence 28, Appr) L	2 F		5	Sequence 26, APPI		1 0	comence 84 Appl	Sequence 2, Appli	Sequence 4, Appri	• (ر د	Sequence b, APP11	Seducince of the	,	Segmence 36, Appl	Sequence 120, APP	7 0	ת	Sequence 65, Appl

ALIGNMENTS

US-09-103-429A-4

Sequence 4, Application US/09103429A Patent No. 6187558

GENERAL INFORMATION:

RESULT

```
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Trichoplusia ni
TISSUE TYPE: peritrophic membrane
US-09-103-429A-4
                                                                                                                                                                                                                                                                            TELEFAX: (607) 256-3628
INFORMATION FOR SEQ ID NO: 4
SEQUENCE CHARACTERISTICS:
LENGTH: 805 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/103,429A
FILING DATE: 24-JUN-1998
CLASSIFICATION: 800
CLASSIFICATION: 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Granados, Robert R
APPLICANT: Wang, Ping
TITLE OF INVENTION: A No. 6187558el Invertebrate Intestinal Mucin
TITLE OF INVENTION: cDNA and Related Products and Methods
NUMBER OF SEQUENCES: 4
                                                                                                                                                                                                                                                                                                                                                               ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A
REGISTRATION NUMBER: 34,390
REFERENCE/DOCKET NUMBER: BTI-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (607) 256-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 118 N
CITY: Ithaca
STATE: NY
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESSEE:
                                                                                                                                                                                                                                                               LENGTH: 805 amino acids TYPE: amino acid
                                                                                                                                                                                                                      STRANDEDNESS:
TOPOLOGY: lir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E: Brown, Pinnisi & Michaels, P.C.
118 No. 6187558th Tioga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    USA
                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                       single
7.7%;
27.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                 BTI-39
  Score
Pred.
  522; DB 4;
No. 6.5e-27;
                     Length 805;
```

Query Match Best Local Similarity

													9											711				
	* . #		***	**************************************					2	*			, f	- -				1		(ME2)		•						
Laboration Coll		ir Fr		3	- S,				-		ja N	4.	1	Seg.			7	3		. 15			* *					
*														4														<i>R</i> .
· -									•																			
2.																												
The state of the s																												
The state of the s		. 1				C.Y																				•		
				~																								
									,																			
\$.			1				- 3		, ÷																			dia.
	- 144 - 14												£			en e												.4
in a sale	e e	44			÷				*	5. 56							•				~	.5	*		٠.	•	-	
	F. 4.	e i			. !	, w	35 · ·		*	.1	+		*4					æ		<i>J</i> .					J . 1	et e	• •	
ind wit		ine collection	•	e p-1		.	127			+.	**						eri						1 - 1		- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	er plant of	62	
	lla		sa sha						44											. 4		a. , , tt				3		
		*			1.5		4					* +	4	 Į2				81		* **						4	* 1	
			*			*	1112			1 [8] A	÷				1.	*		53. 6 3€	1 17 40	-7	*	* 5						
		, p			* · · · · · · · · · · · · · · · · · · ·		5	DW-LST.	4.7		4		*400	**				21. N	7 + **	ar ar			- 2 - 6 - 6		ri Vys			
		*	n	77	ğ.		10	N.								*		50							· · · · · ·	*	i de	
																		÷										
		Α,	÷		5												7											
I .			e"							8																		S. A. C. S.
						•																				, i		and the second of
																												A constant
																												4
					•																							1996
n																								94				28
<u>•</u>																												

```
δ
                                                                                                                                                                                                                                                                                                                                                                                                  Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                             Qy
                                                                                                                                                                                                 Вþ
                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                    Ъ
                                                                                                                                                                                                                                                                                                                                    Дb
                                                                                                                                                                                                                                                                                                                                                                δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   멍
                                                                                             Qy
                                                                                                                                 DЪ
                                                               В
                                                                                                                                                                                                                                                                412 KSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPA-SVKSPSEAKSPAGAKSPAEAKSPVV 470
                                                                                                                                                                                                  645 ---KAKEPPKKVEEEKTPATPKTEVKESKKDEAPKEAQKP-KAEEKEPLTEKPKDSPGEA 700
                                                                                                                                                                                                                      515 APTTPEEPTPTTPEEPAPTTPKAAAAPNTDKEPAPTTPKEPAPTTPKETAPTT 574
                                                                                                                                                                                                                                                                                                                                    529 VKEGAKSLAEAKSPEKAKSPVKEEIKPPAEVKSPEKAKSPMRKEAKSPEKAKTLDVKSPE 588
                                                                                                                                                                                                                                                                                                                                                                404 APTTTKS-----APTTPKEPSPTTTKEPA----PTTPKEPAPTTPKKPAPTTPKEPAPTT 454
                                                                                                                                                                                                                                                                                                                                                                                                     353 PKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKEPA----PTTTKS-----APTTTKEP 403
                                                               745 SEKE----KPK-----KEEVPAAPEKKDTKE-----EKTTESKKREEKPKMEAK 784
                                                                                                                                                                 575 PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAPT 634
                              695 TTPKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTKEPTTIHKSPDE 741
                                                                                             635 TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAP 694
                                                                                                                                 701 KKEEA----KEKKAAAPEEETPAKLGVKEEAKPKEKAEDAKA-----KEPSKP 744
AKEED-----KGLPQEPSKPKTEKAEKSSTDQKDSQPSEKAPED 824
```

Search completed: April 26, 2002, 16:32:37 Job time: 587 sec

```
RESULT ID PLANT IN PROPERTY OF THE PARTY OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ᄝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş
                                                                                                                               Lieberburg I., Spinner N., Snyder S., Anderson J., Goldgaber D., Smulowitz M., Carroll Z., Emanuel B.S., Breitner J., Rubin L.; "Cloning of a cDMA encoding the rat high molecular weight neurofilament peptide (NF-H): developmental and tissue expression the rat, and mapping of its human homologue to chromosomes 1 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NFH_RAT STANDARD; PRT; 831 AA.
P16884; O63368;
01-AUG-1990 (Rel. 15, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT PROTEIN)
((REUROFILAMENT HEAVY POLYPEPTIDE) (NF-H) (FRAGMENT).
                                                                                   -!- FUNCTION
                                                                                                                                                                                                                                                                                  SEQUENCE OF 318-831 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                          Robinson P.A., w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jolles P.;
"The large neurofilament subunit (NF-H)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Breen K.C., Robinson P.A., Wion D., Anderton B.H., "Partial sequence of the rat heavy neurofilament production of putative phosphorylation sites." FEBS Lett. 241:213-218(1988)
                                                                                                                                                                                                                                                            MEDLINE-89184647;
                                                                                                                                                                                                                                                                                                                           FEBS Lett. 209:203-205(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 1-89 AND 243-313 FROM MEDLINE-87080760; PubMed-2878828;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Blochem, Blophys.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chor
Mammalia; Eutheria; Rode
                                                                                                                                                                                                                                                                                                                                                           (NF-H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dautigny
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-88309090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=89065087; PubMed=3143606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1829
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           742
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     508
C. Natl. Acad. Sci. U.S.A. 86:2463-2467(1989).
FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF
NF-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS
SUBSERVED BY THE TWO SMALLER NF DECTRAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SDEE 1969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YSPTSP-TYSPTSPKGSTYSPTSPGY-SPTSP----TYSLTSPAI----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STPE 745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PAPTTPKETAPTTPKEP--APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PTSPTY-SPTTPKY-SPTSP--TYSPTSPVYTPTSPK----YSPTS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT-STTSDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SP-SYSPTSPKYT-----PTSPSYSPSYPEYTPASPK--YSPTSPKYSPTSPK-YS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEP--APT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88309090; PubMed=2457365; A., Pham-Dinh D., Rousse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OF 37-831 FROM
                                                                                                                                                                                                                                                                                                                                                                                                  Wion D., Anderton
                                                                                                                                                                                                                                                                                                                                                                               CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Res.
                                                                                                                                                                                                                                                         PubMed-2928342;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chordata;
Rođentia;
                                                                                                                                                                                                                                                                                                                                                                            for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Commun. 154:1099-1106(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N.A.
                                                                                                                                                                                                                                                                                                                                                                            the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Roussel
                                                                                                                                                                                                                                                                                                                                                                            rat
                                                                                                                                                                                                                                                                                                                                                                                                                                            FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Craniata; Vertebrata; Euteleostomi; Sciurognathi; Muridae; Murinae; Rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Craniata;
                                                                                                                                                                                                                                                                                                                                                                       heavy
                                                                                                                                                                                                                                                                                                                                                                                                в.н.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  c.,
                                                                                                                                                                                                                                                                                                                                                                neurofilament polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Felix J.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the rat:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polypeptide (NF-H)
.";
                   IF PROTEINS: L, N
NEURONAL CALIBER
THAT IS NOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nussbaum J.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cloning and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTYSPTSPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -SPDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rattus
                                                                                                                                                       'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      565
```

218 294 269 269 353

~-PTTPKEPA--

--PTTPKEP--

352

```
DЪ
                                                    Ş
                                                                                              B
                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                   Matches 195;
                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M37227; AAA41693.1; ALT_FRAME.
EMBL; X13804; CAA32038.1; ALT_FRAME.
EMBL; M21964; AAA41695.1; -..
EMBL; J04517; AAA41692.1; -..
PIR; A30796; A30796.
PIR; B25649; A25649.
PIR; B25649; B25649.
PIR; B25649; B25649.
PIR; S02003; S02003.
                                                                                                                                                                                                                                                                               CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P, NFH IS PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF MFH RESULTS IN THE FORMATION OF INTERFLIAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE OF AXONAL CALIBER.
-!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING OF THE LARGER REPORPHIAMENT POLYPEPTIDES (NF M AND NF-H), THE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.
-!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FUNCTION: REE.1 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 783 ONWARD AND IS LONGER DUE TO A FRAMESHIFT.
               243
                                                      162
                                                                                              197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Ew the European Bioinformatics Institute. There are no restruse by non-profit institutions as long as its content modified and this statement is not removed. Usage by and entities requires a license agreement (See http://www.ish-s
                                                                                                                      102 PDTSTTOHNKVSTSPKITTAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTNKQ 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              or send an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Intermediate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -
                                                                                                                                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   $020\u00f3\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u00e4\u
KEAQGEEEEEAEEGGEEAATTS-----PPAEEAASP----EKETKSPVKEEAKSPAEAKS
                                       TSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP----T
                                                                                                                                                                                                                                                                                 482
485
570
591
727
757
769
775
831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   equires a license agreement (S email to license@isb-sib.ch).
                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         276
164
185
193
193
346
373
                                                                                                                                                                                                                                                                        641
164
193
193
193
346
373
482
482
485
771
727
759
769
775
89486
                                                                                                                                                                                            7.3%;
                                                                                                                                                                                                                                                                               MW;
                                                                                                                                                                      59;
                                                                                                                                                             Score 497; DB
Pred. No. 6.1e.
59; Mismatches
                                                                                                                                                                                                                                                                                            51 X 3 AA TANDEM RI

L-> I (IN REF. 2)

I -> S (IN REF. 2)

I -> T (IN REF. 2)

L -> T (IN REF. 2)

M -> T (IN REF. 1)

A -> V (IN REF. 1)

A -> V (IN REF. 2)

P -> S (IN REF. 2)

RK -> KE (IN REF. 2)

RK -> KE (IN REF. 2)

P -> T (IN REF. 2)

A -> V (IN REF. 2)

A -> V (IN REF. 2)

A -> C (IN REF. 2)

A -> C (IN REF. 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  coil;
                                                                                                                                                                                                                                                                           1B0973C3F13EF768
                                                                                 -----KSEKETVIVEEQTEEIQVTEEVTEEED 242
                                                                                                                                                                                     497;
no. 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Neurone; Phosphorylation; Repeat.
                                                                                                                                                                                                                                                                                                 f (IN REF. 4).
N REF. 2).
N REF. 2 AND 4)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      There are no restrictions ong as its content is in oved. Usage by and for cor
                                                                                                                                                                                          DB 1;
.le-14;
                                                                                                                                                                      307;
                                                                                                                                                                                                                                                                                                                                                                  2)
2 A
2 AND
                                                                                                                                                                                                                                                                        2).
2 AND 4).
768 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEATS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                        AND
                                                                                                                                                                                                         Length 831;
                                                                                                                                                                                                                                                                                                                                                                             ND 4).
                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        $ $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       K-S-P
                                                                                                                                                                  146;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     collaboration .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   outstation
                                                                                                                                                               Gaps
                                       217
  293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                               35;
```

MW.

6876FC25692A657E CRC64;

```
_{\rm o}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               뭐
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    γ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ф
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       γQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B
                                                                                                                                                                                                                                                                                                                                                                                                                    БР
                                                                                                                                                                                                                                                                                                                                                                                                                                           Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      В
Вþ
                                                                                                                                                                                                                                                                                                                                                                                             δÃ
                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                          Mus musculus (Mouse).
Eukaryota; Metazoa; Cl
Mammalia; Eutheria; Rc
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1567
                                                                                                                                                                                         01-NOV-1988 (Rel. 09, Created)
01-MAR-1992 (Rel. 21, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT
POLRZA OR RPO2-1 OR RPII215.
                                                                                                                                                                                                                                                                               MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1829
                                                                                                                                                                                                                                                                                                                              1966
                                                                                                                                                                                                                                                                                                                                                                                                                         1876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      216 PTPTTPKEPASTTPKE-PTPTTIKSAPTTPKEPAPTTTKSAPTTPKE-----PAPTTTKE
                                                                                                                                                                                                                                                      RPB1_MOUSE
P08775;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           388
                      SEQUENCE OF 1587-1970 FROM N.A. MEDLINE-86068017; PubMed-2999785;
                                                                       the
                                                                                                          SEQUENCE FROM N.A. MEDLINE=87280135; PubMed=3038894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               566
                                                                                                                                                                                                                                                                                                                                                     742
              Corden J.L.,
                                                                                    "Cloning
                                                                                               Ahearn
 unique structure
                                                           Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTPGSPGSPGPSSPYIPSPGGAMSPSYSPTSPA-YEPRSPGGYTPQSPSYSPTSPSY-SP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEP--APTTPKEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSPMGGISPAMTPWNQGATPAYGAWSPSVGSGMTPGAAGTSPSAASDASGFSPGYSPAWS
                                                             arn J.M. Jr., Bartolomei M.S., West
noing and sequence analysis of the
largest subunit of RNA polymerase
Biol. Chem. 262:10695-10705(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TSPSY-SPTSPNYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-----SPTSPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                PTSPTY-SPTTPKY-SPTSP--TYSPTSPXYTPTSPK-----YSPTS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TPKETAPTTPKGTAPTTLKEPAPTTPK-KPAPKELAPTTTKEPTSTTSDKPAPTTPKGTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPA
                                                                                                                                                                                                                                                                                                                              SDEE
                                                                                                                                                                                                                                                                                                                                                                                                 PAPTTPKETAPTTPKEP--APTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDE 741
                                                                                                                                                                                                                                                                                                                                                                                                                                              PTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPT-STTSDK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SP-SYSPTSPKYT-----PTSPSYSPSYSPEYTPTSPK--YSPTSPKYSPTSPK-YS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PTSP-NYSPTSPSY-SPTSPSY-SPTSP-SYSPSSPR----YTPQSPT-YTPSSPSYSPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTPKEP--APT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -SPTSPSYSPTSPSY-SPTSPSYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSP
                                                                                                                                                                                                                                                                                                                                                                             YSPTSP-TYSPTSPKGSTYSPTSPGY-SPTSP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181;
                                                                                                                                                                                                                                                                                                                                                     STPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity 33.3%; P3
81; Conservative 101;
                                                                                                                                                                                                                                                                                                                                                      745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1970
               Cadena
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA;
                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7.4%;
                                                                                                                                                            Chordata;
Rodentia;
   D.L., A at the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           217205
   Ahearn J.M. Jr., Dahmus M. e carboxyl terminus of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 503.5; DB 1;
Pred. No. 6.8e-14;
                                                                                                                                                             Craniata; Vertebrata; Euteleos o Sciurognathi; Muridae; Murinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                      PRT;
                                                                                                  West
                                                                                                                                                                                                                                                                        1970
                                                                           mouse
II.";
                                                                                                   M.L.,
                                                                                                                                                                         Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                        A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   167;
                                                                                                                                                                                                                                                                                                                                                                                 -TYSLTSPAI ---
                                                                                       genomic
                                                                                                    Cisek
                                                                                                                                                                                                              OH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                         L.J., (
                                                                                                                                                                                                              2.7.7.6) (RPB1).
                   .
E
       largest subunit
                                                                                                                                                                                                                                                                                                                                                                                                                                -PTYSPTSPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1970;
                                                                                                     Corden
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     95;
                                                                                           encoding
                                                                                                                                                                                                                                                                                                                                                                                  -SPDD
                                                                                                                                                                 Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                327
                                                                                                                                                                                                                                                                                                                                                                                   1965
                                                                                                                                                                                                                                                                                                                                                                                                                                 1922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       47;
```

```
Q
                                                                                                                                                                                                                                                                                                                                                               Вp
                                                                                                                                                                                                                                                                                                                                                                                                     QΥ
                                                                                                                             Qy
                                                                                                                                                                             밁
                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                      밁
                                      So
So
                                                                                    멍
밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REVISIONS, AND PRESENCE OF AN ADDITIONAL EXON.

REVISIONS, AND PRESENCE OF AN ADDITIONAL EXON.

MEDLINE=92178992; PubMed=1542581;

Wintzerith M., Acker J., Vicaire S., Vigneron M., Kedinger C.;

"Complete sequence of the human RNA polymerase II largest subuni

"Complete sequence of the human RNA polymerase II largest subuni

"Complete Sequence of the polymerase CATALYZES THE TRANSCR
-1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCR
-1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of eukaryotic RNA polymerase II."; proc. Natl. Acad. Sci. U.S.A. 82:7934-7938(1985).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the EM the European Bioinformatics Institute. There are no restruse by non-profit institutions as long as its content use by non-profit institutions as long as its content modified and this statement is not removed. Usage by and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00623; RNA_PO1_A; 1.
Pfam; PF01854; RNA_PO1_A2; 1.
PROSITE; PS00115; RNA_PO1_II_REPEAT; 42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              entities requires a license agreement (See http://www.lisb-slb.or send an email to license@isb-slb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR000684; RNA_polII_repeat InterPro; IPR000722; RNA_pol_A. InterPro; IPR002879; RNA_pol_A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                 1507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Transferase;
                                                                                                                                                                               1625
                                                                                                                                                                                                                                                                          1567
                                                                                                                                                                                                                                                                                                                                                                                                               216
                                                                                                                                                                                                                             328
                                                                                                                                                                                                                                                                                                                   270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUBUNIT: RNA POLYMERASE II CONSISTS OF SUBCELLULAR LOCATION: NUCLEAR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                III FOR 5S AND TRNA GENES. SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              THE PHOSPHORYLATION ACTIVATES POL2.

MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES EQUID IN EUKARYOTIC NUCLET: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE PROCURSOR, AND POLYMERASE PRECURSOR, AND POLYMERASE PRECURSOR, AND POLYMERASE PROCURSOR, AND POLYMERASE PROCURSOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PTM: THE TANDEM 7 RESIDUES REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBSTRATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MGI:98086; Rpo2-1
                                                                                                                                                                                                                                                                        PTPGSPGSPGPSSPYIPSPGGAMSPSYSPTSPA-YEPRSPGGYTPQSPSYSPTSPSY-SP 1624
                                                                                                                                                                                                                                                                                                                                                               PSPMGGISPAMTPWNQGATPAYGAWSPSVGSGMTPGAAGFSPSAASDASGFSPGYSPAWS 1566
                                                                                                                                                                                                                                                                                                                                                                                                          PTPTTPKEPASTTPKE-PTPTTIKSAPTTPKEPAPTTTKSAPTTPKE-----PAPTTTKE 269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A28490; A28490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             M12130; AAA40071.1;
M14101; AAA40071.1;
SY-SPTSPSY-SPTSPSY-SPTSPNY-SPTSPNYT-PTSP-SYSPTSP-SYSPTSPNY-T 1777
                                                                                                                                                                                                                        TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTTKEPSPTTPKE 387
                                                                                                                                                                                                                                                                                                                 PAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTTPKKPAPTTPKEP--APTTPKEPTP
                                        KEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPA
                                                                                                                               PAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTP 447
                                                                                                                                                                               TSPSY-SPTSPNYSPTSPSY-SPTSPSY-SPTSPSY----
                                                                                      SPTSPSYSPTSPSY-SPTSPSYSPTSPSY-SPTSPSY-SPTSPSY-SPTSPSY-SPTSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1590
1498
1499
1970
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
Nuclear protein; Phosphorylation; Zinc-finger.
Nuclear protein; Phosphorylation; Zinc-finger.
R7 C2H2-TYPE (POTENTIAL);
S90 1958 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
498 1498 P -> R (IN REF. 1 AND 2).
499 1536 MISSING (IN REF. 1 AND 2).
490 1536 MISSING (IN REF. 1 AND 2).
497 AA; 217175 MW; 7D76F38FD92A657E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 498.5; DB 1
Pred. No. 1.1e-13;
1; Mismatches 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAN BE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12 DIFFERENT SUBUNITS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HIGHLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORYLATED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLYMERASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        subunit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a collaboration
                                                                                                                                                                                                SPTSPSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  η
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                       1725
                                                                                                                                                                                                1671
                                                                                                                                                                                                                                                                                                                                    327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  og
                                                            507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           47;
```

```
В
                                  S
                                                    밁
                                                                Š
                                                                                      밁
                                                                                                   Ş
                                                                                                                     밁
                                                                                                                                     Ş
                                                                                                                                                          밁
                                                                                                                                                                       Ş
                                                                                                                                                                                           В
                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                밁
                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                      ₽
                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                          문
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Sim
Matches 223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
VARIANT
VARIANT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
                                                     464
                                                                                       437
                                                                                                                         385
                                                                                                                                                           329
                                                                                                                                                                                                             560
                                                                                                                                                                                                                               271
                                                                                                                                                                                                                                                                                   440
                                                                                                                                                                                                                                                                                                    219
                                                                                                                                                                                                                                                                                                                    380
                                                                                                                                                                                                                                                                                                                                     164
                                                                                                                                                                                                                                                                                                                                                                       104
                                                                                                                                                                                                                                                                                                                                                                                                                                                            157
                                                                                                                                                                                                                                                                                                                                                                                                           52
                                                                                                                                                                                                                                                                                                                                                                                                                                           ω
MPRVRKPKTTPTPR 919
                      KVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPK----T
                                                             LRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAP
                                                                                          KEPTTIHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERD
                                                  -TTTPTTTTTP----TTTTTKAT----
                                                                                                                    ISGTKWCFYSTSQVAA - - TKTTTTPTTTTTTTTTTTTTKATTTTP - - - - -
                                                                                                                             TKGCCFDSSIPQTKWCFYTLSQVADCKVEPSQRVDCGFRGIT----ADQCRQKNCCFDSS
                                                                                                                                                                                                                                                                       KKPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKETAPTTPKKLTPTTPEKLA
                                                                                                                                                                                                       KEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTT
                                                                                                                                                                                                                                          PTTPEKPAPTTPEELAPTTPEEPTPTTPEEPAPTTPKAAAPNTPKEPAPTTPKEPAPTTP
                                                                                                                                                                                                                                                                                                       VKPSKREMCGSKGITKKQ-----CKKKNCCFDPKGHGGIHCFHRKPKGHSHEEHTTTTK-
                                                                                                                                                                                                                                                                                                                                                                GKAPATAAAPVPTTAASKAPTTAAAATHSTAAAAAAPTTAASAAKSKERSTSSSSEEEHCH
                                                                                                                                                                                                                                                                                                                                                                                 EPAPTTTKEPAPTTTKS-APTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPA-
                                                                                                                                                                                                                                                                                                                                                                                                                   TPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                     TTAAAVAATGKDTTAAAEGSAAAEKTAA----AGEVSAPPT--AAVAATGEDATT----
                                                                                                                                                                                                                                                                                                                                                                                                                                              TINKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEP
                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       374
526
536
5746
5743
583
583
593
6821
6821
6741
276
415
                                                                                                                                                                    -TAPTT-----PKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                   AAATAAAETTAAAGEAPTTTTAPATTAAGKAPTTAAATAPTTAAAGAPTTAT
                                                                                                                                                                                                                                                                                                 APTTIQIATTTT---TPTTT----
                                                                                                                                                                                      TTTTTKATTTTTTSGECKMEPSK------REDCGYSGITESQCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ą.
                                                                                                                                                                                                                                                             TTTTKATPTTT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                391
552
553
599
598
598
598
598
598
547
546
546
558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        67774
                                                                                -TTKA-----KAT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ₩,
             PSKRADCGYPGITESQCRSKGCCFDSSIPQTKWCFYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 509.5;
Pred. No. 1.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Y SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F085277F1ED2FD40 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                               TTTPTTTTTTTTTTKATTTTTTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5; DB 1;
1.6e-14;
hes 276;
                                                                                                                                                                                                                                                             TTTKATTTTT---
                                                                                                                                                                                                                                                                                               -TTTTKATPTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                    -TTTTTTPTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        247;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
             566
                             905
                                                               849
                                                                                463
                                                                                                 789
                                                                                                                                   729
                                                                                                                                                    384
                                                                                                                                                                     669
                                                                                                                                                                                                        619
                                                                                                                                                                                                                                                                             499
                                                                                                                                                                                                                                                                                                               439
                                                                                                                                                                                                                         289
                                                                                                                                                                                                                                                           270
                                                                                                                                                                                                                                                                                                                                218
                                                                                                                                                                                                                                                                                                                                                                 163
                                                                                                                                                                                                                                                                                                                                                                                                                     276
                                                                                                                                                                                                                                                                                                                                                                                                                                     51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30;
```

```
RESULT 13
RPBILHUMAN
ID _RPBI_H
AD _PUMAN
ID _PAUGE
DE DNA-DI
GN POLR2A
OS HOMO S
OC MAMMAI
OC MAMMAI
OC MAMMAI
OC MAMMAI
RP SEQUEN
RP SEQUEN
RP SEQUEN
RY MEDLIN
RA MICHA
RY MICHIA
RY MEDLIN
RA MICHA
RY MEDLIN
RA MICHA
RY MEDLIN
RA MICHA
RY FORD
CC MAMMAI
OC -I- SU
CC -I- SU

R InterPro; IPR000684; RNA_POLIT_repeat.

R InterPro; IPR000722; RNA_POL_A.

R InterPro; IPR000723; RNA_POL_A.

R InterPro; IPR002879; RNA_POL_A.

R Pfam; PF01854; RNA_POL_A.

R Pfam; PF01854; RNA_POL_II_REPEAT; 43.

R PROSITE; PS00115; RNA_POL_II_REPEAT; 43.

R PROSITE; PS00115; RNA_POL_II_REPEAT; 43.

R PROSITE; PS00115; RNA_POL_II_REPEAT; 43.

R PGam; PF01854; RNA_POL_A; 1.

R Pfam; Pf01854; RNA_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X63564; CAA45125.1; -
EMBL; X74874; CAA52862.1; J
EMBL; X74873; CAA52862.1; J
EMBL; X74872; CAA52862.1; J
EMBL; X74871; CAA52862.1; J
EMBL; X74870; CAA52862.1; J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collable between the Swiss Institute of Bioinformatics and the EMBL outsit the European Bioinformatics Institute. There are no restrictions use by non-profit institutions as long as its content is in modified and this statement is not removed. Usage by and for contential the statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RPB1_HUMAN STAN
P24928;
01-MAR-1992 (Rel. 2
01-MAR-1992 (Rel. 2
20-AUG-2001 (Rel. 4
DNA-DIRECTED RNA PC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          entities requires a license agreement (some send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE=95347616; PubMed=7622068;
Mita K., Tsuji H., Morimyo M., Ta
Tichimura S., Yamauchi M., Hongo E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDITINE-92178992; PubMed=1542581; Wintzerith M., Acker J., Vicaire S., "Complete sequence of the human RNA Nucleic Acids Res. 20:910-910(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=9606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLR2A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
SUBCELLULAR LOCATION: NUCLEAR.
PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
THE PHOSPHORYLATION ACTIVATES POLZ.
MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES AR
FOUND IN LUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND TENA GENES.
III FOR 5S AND TENA GENES.
SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       b human gene encoding the largest subunit of RNA polymerase 1 = 159:285-286(1995).

FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCHOF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBSTRATES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         :|:| | |: |
LPQVADCKVAPSSR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ji H., Morimyo M., Takahashi E., Nenol
Yamauchi M., Hongo E., Hayashi A.;
eneocoding the largest subunit of RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACTIVITY: N NUCLEOSIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               L. 21, Created)
L. 21, Last sequ
L. 40, Last anno
A POLYMERASE II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOINED JOINED JOINED JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sequence update)
annotation update)
E II LARGEST SUBUNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Craniata; Vertebrata; Catarrhini; Hominidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polymerase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRIPHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Vigneron M., Kedinger C.;
polymerase II largest subunit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 noved. Usage by and for (See http://www.isb-sib
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (E)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ı
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.7.7.6) (RPB1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N PYROPHOSPHATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           collaboration -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ch/announce,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   II.";
```

δÃ

δÃ 밁

δ 망

B

δÃ

δõ

В

망

```
Hauser F., Hoffmann W.;
"P-domains as shuffled cysteine-rich modules in integumentary C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic polymorphism.";
J. Biol. Chem. 267:24600-0007
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               237
                                                                                                                                                                                        XENLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   186
                                                                                                                                                                                                                                                                                                                                                                                                                              630
                                                                                                                                                                                                                                                                                                                                                                         629
                                                                                                                                                                                                                                                                                                                                                                                                             584
                                                                                                                                                                                                                           828
                                                                                                                                                                                                                                                                                                                    791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAPTTPKEPAPTT----PKEPTPTT--PKEPAPTTKEPAPTTPKEPAPT-APKKPAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SYPSIYKPKASYVSSYKSKKTYPPTYKPKISYPPTYKPKPSYP----PTYKPKVTYPPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RKPSYTPYKPKATYPPTYKPKITYP----PTYKRKPSYT----PYKPKTTYPPTYKPKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYKPKITYPPTYKPKPSYPPSYKPKKTYSPTYKPKITYPPTYKPKPSYPPSYKPKTTYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            K-PKPSYPPTYKPKITYPPT-YKPKPSYPTPYKQKPSYPPIYKSKSSYPTSYKSKKTYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEPAPTTPKEPAPTTTKEPSPTTPKEPAPTTTKSAPT----TTKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      T-YKPKISYPPTYKPKASYVSSYKSKKTYPPTYKPKISYPPTYKPKPSYPPTYKPKITYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APTAPKEPAPTTPKETAP-----TTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EPSPTTTKEPA-PTTPKEPAPTTPKKP-APT-TPK-----EPAPTTPKEPAPTTTKKP
                                                                                                                                                                                                                                                                                                                                                                                                                                              PTYKPKITYPPTYK-----PKPSYPPSYKPKTTYPPTYKPKIRYPPTYKPKASYPPTYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                TT--PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPT-TPKGTAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTYKPKPSYPPTYKRKITYPPTYKRKPSYPTPYKQKPSYPPIYKSKSSYPTSYKSKKTYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PT-PTTPEEPAPTTPKAAAPNTPK-EPAPTTPFEBPAPTTP-----KEPAPTT--PKETAP
                                                                                                                                                                                                                                                                                                                                                                                         PTT--PKETAPTTPKEPAPTTPK--KPAPTTPET-----PPPTTSEVSTPTTTKEPTT
                                                                                                                                                                                                                                                                                                                                                                                                                             EPAPTTPKEPAPTTPKGTAPTTLKEPAPTTP----KKPAPKELAPTTTKGPTSTTSDKPA
                                                                                                                                                                                                                                            MPRVRKPKTT--PTPRKMTSTMPELNPTSRIAEAMLQTTTRPNQ
                                                                                                                                                                                                                                                                          VTTTKKTITTEIMNKPEETAKPKDRATNSKATTPKPQ---KPTKAPKKP-TSTKKPK-T;
                                                                                                                                                                                                                                                                                                  ISYPSTYKAKP-----SYPPTYKPKPSYASSYKPKIRYPPTYKPKPSYASSYK-----
                                                                                                                                                                                                                                                                                                                  RTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPK
                                                                                                                                                                                                                                                                                                                                                     IHKSPDESTPELSAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKT----TERDL
                                                                                                                                                                                                                            YPPTYKPKISYPPTYKPKITYPPTYKPKISYPPAYKPKISYPSQ
                                                                                                                                                                                                                                                                                                                                                                          PPTYKPKITYPPTYK-PKPSYPPSYRPKITYPPTYKPKKSYPQAYKSKGSYPPSYQPKKT
                                                                                                                                                                                                                                                                                                                                     PPSYKPKKTYPPTYK----PKISYPPTYK----TKPSYPASYKRKTSYPPTYKPK
                                                                                                                                                                                                                                                                                                                                                                                                             ----PKITYPPTYK-PKPSYPTPYKQKPSYPPIYKSKSYPTAYKSKKTY
                                                                                                                                                                                                                                                               -RYPPTYKPKP---
                                                                                                                                                                                                                                                                -SYASSYKPKIRYPPTYKPKPSYASSYKPKIT
                                                                                                                                                                                                                              871
                                                                                                                                                                                                                                                 947
                                  mucin
                                                                                                                                                                                                                                                                                                       PK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                           735
                                                                                                                                                                                                                                                                                                                                                                                                                 628
                                                                                                                                                                                                                                                                                                                                                                                                                                   685
                                                                                                                                                                                                                                                                                                                                                                                                                                                     583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                350
                                                                                                                                                                                                                                                                                                                                                                             687
                                                                                                                                                                                                                                                                                     905
                                                                                                                                                                                                                                                                                                       785
                                                                                                                                                                                                                                                                    827
```

멍

QΥ В

δÃ

В

밁 δÃ QY В Qγ Вр Qy В δÃ Ъ

```
the European Bioinformatics Institute. Ther use by non-profit institutions as long a modified and this statement is not removed. entities requires a license agreement (See For send an email to license@isb-sib.ch).
                                                                                                                                                                                                                              This SWI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -
EMBL; L02115; AAA74725.1;
PIR; A45155; A45155.
HSSP; P04002; 1WFA.
                                                                                                                                                                                                                                                                                                                                  SUBCELLULAR LOCATION: SECRETED.
ALTERNATIVE PRODUCTS: A NUMBER OF DIFFERENT
MAY BE PRODUCED BY ALTERNATIVE SPLICING.
TISSUE SPECIFICITY: SKIN.
PIM: EXTENSIVELY O-GLYCOSYLATED.
                                                                                                                                                                                                                                                                                                            SIMILARITY: CONTAINS 6 P-TYPE (TREFOIL) DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT
                                                                                                                                                                                                                              SWISS-PROT entry is copyright. It is produced through a collaboration een the Swiss Institute of Bioinformatics and the EMBL outstation
                                                                                                                                  tatement is not removed. Usage by and for commercial license agreement (See http://www.isb-sib.ch/announce
                                                                                                                                                                                     ormatics and the EMBL outst There are no restrictions ong as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                           FORMS
                                                                                                                                                                                                                                                                                                                                                                                                                           OF,
                                                                                                                                                                                                                                                                                                                                                                                                                              THE PROTEIN
                                                                                                                                                       is for
                                                                                                                                                                                                                        on
                                                                                                                                                                                                   no
```

Repeat; REPEAT REPEAT REPEAT REPEAT REPEAT REPEAT REPEAT DOMAIN DOMAIN REPEAT DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID
DISULFID REPEAT REPEAT REPEAT DOMAIN DOMAIN DOMAIN DOMAIN DOMAIN REPEAT REPEAT REPEAT REPEAT REPEAT PROSITE; DOMAIN REPEAT DOMAIN DISULFID Amphibian skin; PS00025; P. 137 161 218 218 225 225 240 250 260 276 276 288 295 288 295 306 306 81 97 105 113 144 411 419 431 443 450 460 472 491 491 491 491 1188 515 522 523 531 661 1188 661 1187 332 332 332 332 332 TREFOIL; 8 X 8 AA AF.
ALA/THR-RICH.
1-1.
1-2.
1-3.
1-4.
20 1-5.
1-6.
1-7. Glycoprotein; Alternative splicing 3-8.
3-10.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3-11.
3 1-8.
p-TYPE
8 X APF
2-1.
2-2.
2-3.
2-4.
2-5.
2-6.
2-7. 3-4 2-8.
2-8.
2-TYPE 2.
P-TYPE 3.
12 x APPROXIMATE TANDEM REPEATS.
THR-RICH. -TYPE 4.
-TYPE 5.
-TYPE 6.
-TYPE 6.
Y SIMILARITY.
Y SIMILARITY. APPROXIMATE TANDEM REPEATS, THR-RICH. AA APPROXIMATE TANDEM REPEATS,

```
, <sup>B</sup>
                                     ş
                                                               8
                                                                                      ð
                                                                                                                  밁
                                                                                                                                      Š
                                                                                                                                                                    В
                                                                                                                                                                                           õ
                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                         밁
                                                                                                                                                                                                                                                                                              á
                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                                                                                                                                  밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SOFFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 7.8%;
Best Local Similarity 26.2%;
Matches 233; Conservative
            814
                                                                                                                                            805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                 697
                                                                                                                                                                                                                                                                                                    651
                                                                                                                                                                                                                                                                                                                              470
                                                                                                                                                                                                                                                                                                                                                       599
                                                                                                                                                                                                                                                                                                                                                                                414
                                                                                                                                                                                                                                                                                                                                                                                                                                                           494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               244 PKEPAPTTTKSAPTTP-KEPAPTTTKEPAPTTPKEPAPTTTKEPAPT----TTKSAPTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  188 APTSKVLAKP-TPK---AETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   æ
                           PRKMTSTMPELNPTSRIAEAM----LQTTTRPNQTPNSKLVEVN-----PKS
    LRDLQTTDVSLLAIAATLDAIGEKLKDQKARNQQVMDRLCEIEKILGPPKS
                                                      EVPITAGDNPDNTSVGISEVVPTIAEKPVEEVPTS--EIPEQSSSPS--DSVPVAKITPL
                                                                             ITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                                        DIAIPVIDPPVPQEIAVAEIPETDTKPAEVIVEQSTIPIEAPVPEVSKYAEPVISEAPAA
                                                                                                                                 KETATTTEKTTESKITATTTQVTST-----TTQDTTPFKITTLKTTTLA-PKVTTTKKT
                                                                                                                                                          : | |:|:
TVE-FPEAVAEKVLDPAI--TEAPVTTQEPDVANINDGAPATEI-TTPAVEIVTAAAEVS
                                                                                                                                                                                SAEPTPKALENSPKEPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTP--ETTTAAPKMT
                                                                                                                                                                                                             EVAVAPITAPEPIP----EPEPSLATPTEPIPVEAPVVIQEAVDAVEVPVTETSTSIPET
                                                                                                                                                                                                                                      PKEPAPTTPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPD-----ESTPEL
                                                                                                                                                                                                                                                                 APPEAAADLIIEPVEPPAPIPDLLEQTTSVPAVEAAESTSS--PIPETSLPPPNEAVASP
                                                                                                                                                                                                                                                                                        TLKEPA-----PTTPKKPAPKELAPTTT----KGPTSTTSDKPAPTT----PKETAPTT
                                                                                                                                                                                                                                                                                                                                            LAPTTTKEPTSTTSDKPAPTTPKGTAPTTP-----KEPAPTTPKEP--APTTPKGTAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAPIVSTPP-
                                                                                                                                                                                                                                                                                                                                                                     V----LPPVAAEPVPAVVAEETPETPAPASAPVTTAALDIPEVAPVTAAPSDAPAEAPSA
                                                                                                                                                                                                                                                                                                                                                                                              TPKEPAPTTPKEPAPTTPKETAPTTP-KGTAPTTLKE-----PAPTTPKKPAPKE
                                                                                                                                                                                                                                                                                                                                                                                                                       AAESIPAPVVATTPVPAT---LAVTDPDVTASAVPELPPVIAPSPVPSAVAETPVDLAPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                TPEKL-APTTPEKPAPTTPEELAPTTPEEPTPTTPEEP---APTTPKAAAPNTPKEPAPT 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ETPALAPVVAESQVAANTVVATPPTPAPEPETIAPPVVAETPEVASVAVAETTPPVVPPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PLAAAEPVVVAPPATETPVVAPAAASPHVSVAPAVETAVVAPVSASTEPPVAAATLTTAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PTTTKEP---APTTTKSAPTTPKEPSPTTTKEPA-PTTPKEPAPTTPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAVPAAVPVVAPVLAPAVAPAVAPVVAETPAPPPVAEIPVATIPECVAPLIPEVSVVATK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PKEPA--PTAPKKPAPTTPKEPAPTTPKEPAPTTTKE-PSPTTPKEPAP----TTTKSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PVAATPTPVVQIPVAAPVIATPPVAASAPTPAAVTPVISPVIASPVVPANTTVPVAAPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KEPAPTTPKKPAP-----TTPKEPAPTTPKEPTP-----TTPKEPAPTTKEPAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VTIPAPAPIAAASVTPVASVAPPVVAAPTP----PAASPVSTPVAVAQIPVAVSAPVAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SPVSAPVAAPVTPSAVAAPVQVVSPAAVAP----APAAPIAVTPVAPPPTLASVQPAT
                                                                                                                                                                                                                                                                                                                   -TTASVPETTAPPAAVPTEPIDVSVLSEAAIETPVAPPVEVTTEVAVADV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   534
699
703
721
84781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ₹
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 530;
Pred. No. 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              D<H>>>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A -> E (IN REF. 2).
I -> T (IN REF. 2).
V -> L (IN REF. 2).
D -> E (IN REF. 2).
2110417E0B0E7CFE CI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
.9e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                       813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   493
                                                                                                         757
                                                                                                                                    857
                                                                                                                                                                                    804
                                                                                                                                                                                                               641
                                                                                                                                                                                                                                                                 585
                                                                                                                                                                                                                                                                                          696
                                                                                                                                                                                                                                                                                                                                             650
                                                                                                                                                                                                                                                                                                                                                                        469
                                                                                                                                                                                                                                                                                                                                                                                                                        413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41;
```

RESULT

```
FPI_MYTCO
AC Q2543
AC Q2643
AC
                                     Š
                                                                                   В
                                                                                                                                Ş
                                                                                                                                                                                   밁
                                                                                                                                                                                                                           Ş
                                                                                                                                                                                                                                                                             Ъ
                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Sim
Matches 276;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRINTS;
Signal;
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Its evolutionary implications.";

J. Mol. Evol. 43:348-356(1996).

-I- FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSEL'S PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

-I- SUBCELLULAR LOCATION: SECRETED.

-I- TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

-I- DOMAIN: ALMOST EXCLUSIVELY COMPOSED OF REPEATS OF A DECAPEPTIDE.

-I- PTM: PROLINES IN THE REPEAT ARE HYDROXYATED (SINGLE OR DI-) AND ALSO TYROSINE (THUS PRODUCING DOPA = 3,4-DIHYDROXYPHENYLALANINE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002964; Adhesive_plaq.
InterPro; IPR002965; P_rich_extensn.
PRINTS; PR01216; ADHESIVEI.
PRINTS; PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a c between the Swiss Institute of Bioinformatics and the EMBL the European Bioinformatics Institute. There are no restrict use by non-profit institutions as its content is modified and this statement is not removed. Usage by and fo entitles requires a license agreement (See http://www.isb-sib.or send an email to license@isb-sib.ch).
254 SAPT-TPKEPAPT--TTKEPAPTTPKEPAPTTTKEPAPTTTKSAPTTPKEPAPTT--PKK 308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D63777; BAA09850.1;
                                                                                     140
                                                                                                                                     195
                                                                                                                                                                                                                                135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FP1_MYTCO
Q25434;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Inoue K., T
Harayama S.
                                                                                                                                                                                   85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-96394686; PubMed-8798340;
Inoue K., Takeuchi Y., Takeyama S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Foot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Mollusca; B
Mytiloidea; Mytilidae; Mytilus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Myttlus coruscus (Sea mussel).
Myttlus coruscus (Sea mussel).
Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence up
20-AUG-2001 (Rel. 40, Last annotation
ADHESIVE PLAQUE MATRIX PROTEIN PRECURS
                                                                                                                                                                                                                                                                                                                           84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Adhesive protein cDNA sequence ts evolutionary implications.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-42192,
                                                                                                                                                                                                                                                                             29
                                                                    YPPTYK-----PKITYP--PTYKQKPSYPPSYKPKTTYPPTYK-----PKITYPPTYK 185
                                                                                                                                                                                                                                                                                                       VVDEAGSGLDNGDFKVTTPDTST--TQH-----NKVSTSPKITTAKPINPRPSLPPNS
                                                                                                              AKPTPKAETTTKGPALTTPKEPT-PTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTK 253
                                                                                                                                                                                                            DTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVL 194
                                                                                                                                                                 YPTKRYQPTYGSKTNYPPIYKPIAKKLSSYKAIKTTYPAYKAKTSYPPSYK-----HKIT
                                                                                                                                                                                                                                                            VYGSAYSGASAGAYK-TLPGSHPYGSKHVPVYKPMNKIPT-PYI--SKKSYPAPYKPKGY
                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          872 AA;
                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   184
213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1
21
21
124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydroxylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       192
221
101677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  872
41
872
                                                                                                                                                                                                                                                                                                                                                                                          7.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
ADHESIVE LAQUE MATRIX PROTEIN.
NONREPETIVE LINKER,
TANDEM REPEATS OF Y-K-[PS]-K-[IP]-{ST]-Y-P-[PST]-[ST].
NONAPETIDE 1.
NONAPETIDE 2.
NONAPETIDE 2.
MW; 98CC70D7C75FF3C4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                   ; 08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       M.
                                                                                                                                                                                                                                                                                                                                                         Score 518.5; DB 1
Pred. No. 8.6e-15;
0; Mismatches 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Otation update)
PRECURSOR (FOOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of the mussel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yamaha E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A
                                                                                                                                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                 407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mytilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Yamazaki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEIN 1) (MCFP1).
                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             coruscus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F)
                                                                                                                                                                                                                                                                                                                                                                                                      872;
                                                                                                                                                                                                                                                                                                                                                           181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a collaboration MBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Odo
                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    commercia;
                                                                                                                                                                 139
                                                                                                                                                                                                                                                              84
                                                                                                                                                                                                                                                                                                         134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                         54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        way
```

```
Qy
                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                    δÃ
                                                                                                                                                                                                                                                                                                                                                                              Вþ
                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                              DЪ
                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                       맒
                                              ф
                                                                       δõ
                                                                                                     В
                                                                                                                          δÃ
                                                                                                                                                        밁
                                                                                                                                                                              γQ
               Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
Matches 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SW:
between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   modified and this statement is not removed entities requires a license agreement (See or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             use
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M61732; AAA30255.1;
PIR; JH0557; JH0557.
HSSP; P29768; 1DIL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002860; 1
Pfam; PF02012; BNR; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Phosphorylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hydrolase; Glycosidase; Glycoprotein; Repeat; GPI-anchor;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                           217
                                                                                                                                                                                                                                                                                                                                                                                  599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ween the Swiss Institute of Bioinforma
European Bioinformatics Institute. The
by non-profit institutions as long
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEVELOPMENTAL STAGE: MAXIMAL ACTIVITY IN TRYPOMASI IN EPIMASTIGOTES AND NO DETECTABLE IN AMASTIGOTES. MISCELLANEOUS: THE VARIABLE LENGTHS OF THE LONG TA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN COULD ACCOUNT IN PART FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POSSIBLE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SWISS-PROT entry is copyright. It is produced through a collaborative een the Swiss Institute of Bioinformatics and the EMBL outstation European Bioinformatics Institute. There are no restrictions on in
                                                                                                                                                                                                                                                                                                                                                                                                         TPTTPKE-PASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTP
                                                                                                                                                                                                                                                                                                 KEPTP-TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAPTTPKEPAPTTKEPS
                                                                                                                                                                                                                                                                                                                            \verb|ADSSAHGTPSTPVDSSAHSTPSTPVDSSAHGAPSTPADSSAHGTPSTP|\\
                                                                                                                                                                                                                                                                                                                                                     KE-PAPTTTKEPAPTTTKSAPTTPKE-PAPTTPKKP----APTTPKEPA----PTTP 322
                                                                                                                                                                                                                                                                                                                                                                                 TPSTPADSSAHSTPSTPVDS----SAHSTPSTPADSSAHGTPSTPVDSSAHGT----PSTP
                                                                                                                                                                                                                  HGTPSTPADSSAHSTPST---PADSSAHGTPSTPVDSSAHST----PSTPVDSSAHGTPS
                                                                                                                                                                                                                                        PTTPKEPAPTTTKSAPTTTKEPAPTTTKSAPTTPKEPSPTTTKEPAPTTPKEPA----PT
                                                                                                                                                                                                                                                                       VDSSAHGTPSTPADSSAHSTPSTPADSS-----AHSTPSTPADSSAHSTPSTPVDSSA
                                                                          PKEPAPTTPKEPAPTTPKEPAPTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPT
                                                                                                       STPVDSSAHSTPSTPADSSAHS-TPSTPVDSSAHSTPSTPADSSAHGTPSTPVDSSAHGT
                                                                                                                               TTP-EKLAPTTPEKPAPTTPEELAPTTPEE----PTPTTPEEPA----PTTP-KAAAPNT
                                                                                                                                                            TPVDSSAHSTPSTPVDSSAHGTPSTPVDSSAHSTPSTPADSSAHST-PSTPADSSAHGTP
                                                                                                                                                                                        TP-KKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKL---
                     TTKEPTSTTSDKPAPTTPKGTAPTTPKE-PAPTTPKEPAPTTPKGTAPTTLKEPAPTTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
163
209
458
589
342
342
394
1125
1162
                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BELONGS TO FAMILY 33 OF CONTAINS 3 BNR REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   457
34
174
220
588
1120
                                                    AHSTPSTPADSSAHST-PSTPADSS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                        29.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATTACHED
                                                                                                                                                                                                                                                                                                                                                                                                                                         100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FAMILY 33 OF GLYCOSYL HYDROLASES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FIBRONECTIN TYPE-III.

44 X 12 AA TANDEM REPEATS, LTR DOMA
N-LINKED (GLCNAC. . .) (POTENTIAL).
MW; 07049221897C6A40 CRC64;
PSTPADSSAHSTPSTPADSSAHSTPSTPVDSSAHSTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 530.
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BNR
BNR
BNR
                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             THS OF THE LONG TANDEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  THE MEMBRANE BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                         .6e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   as its content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IN TRYPOMASTIGOTES,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Usage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEATS, LTR DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Уď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GPI-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MINIMUM
                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                               764
                                                                                                                                                                                                                                                                                                           381
                                                                                                                                                                                                                           817
    1033
                                                                                                                                          542
                                                                                                                                                                     876
                                                                                                                                                                                               492
                                                                                    602
                                                                                                               935
```

```
RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δÃ
 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CPN_DROME
Q02910;
Q1-OCT-1993
Calcium-binding.
CONFLICT 36
CONFLICT 43
CONFLICT 76
CONFLICT 76
CONFLICT 100
CONFLICT 126
CONFLICT 154
CONFLICT 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1034
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
STRAIN=CANTON-S;
MEDLINE=93165729; PubMed=8094559;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pterygota; Ne Ephydroidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Drosophila
Eukaryota;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CALPHOTIN.
CPN OR CAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-OCT-1993
01-FEB-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                Proc. Natl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 662
                                                                                                                                                               the
                                                                                                                                                                                                                                      This
                                                                                                                                                                                                                                                                                                                                         calcium and contains a leucine zipper.";
proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).
-i- FUNCTION: MIGHT FUNCTION AS A CALCIUM-SEQUESTERING "
REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT
                                                                                                                                                                                                                                                                                                                                                                                      Ballinger D.G., Xue N., Harshman K.D.; "A Drosophila photoreceptor cell-specific protein, calphotin,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   718
                                                                                                                             EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=93165730; PubMed=8434015;
                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-CANTON-S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       777
                                                                                                                    PIR;
                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "Calphotin:
                                                                                                                               EMBL; L02111; AAA28405.1; EMBL; L05080; AAA28420.1;
                                                                                                                                                                                                                                                                                                        OF CA+2 PER MOL OF PROTEIN.
SUBUNIT: HOMODIMER (PROBABLE).
SUBCELLULAR LOCATION: CYTOPLASMIC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KPAPKELAPTTTKGPTSTTSDKPAPTTPK----ETAPTTPKEPAPTTPKKPAPTTPETPP
                                                                                                                                                                                                                                                                       TISSUE SPECIFICITY: SOMP COMPOUND EYES AND OCELLI DEVELOPMENTAL STAGE: EXP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ΜŢ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTTSEVSTPTTTKEPTTIHKSPDESTP-ELSAEPTPKALENSPKEPGVPTTKTPAATKPE 776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TP----ADSSAHGTPSTPADSSAHSTPSTPVDSSAHSTPSTPADSS
                                                                                                                                                                                                              European
                                                                                                                                                                                                                                                            DEVELOPMENT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GT 1127
                                                                                                                                                                                                           SWISS-PROT entry is copyright. It is produced through a collaboration sen the Swiss Institute of Bioinformatics and the EMBL outstation in Suropean Bioinformatics Institute. There are no restrictions on its suropean Bioinformatics Institute.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           J.H., Benzer S., Rudnicka M., Miller C.A.; otin: a Drosophila photoreceptor cell calcium-binding Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       778
                                                                                                          FBgn0010218; Cpn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     melanogaster (Fruit fly).
Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
a; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Rel.
(Rel.
(Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27, Created)27, Last sequence update)28, Last annotation update)
  36
43
64
76
100
127
154
                                                                                                                                                                                                                                                                                               SOMA AND AXONS
                                                                                                                                                                                                                                                                          EXPRESSED EARLY IN
 A -> AVAPAVVA (IN REF. 1) T (IN REF. 2).
I -> Y (IN REF. 2).
I -> V (IN REF. 2).
T -> A (IN REF. 2).
P -> AP (IN REF. 2).
VQ -> AP (IN REF. 2).
I -> V (IN REF. 2).
S -> T (IN REF. 2).
                                                                                                                                                                                                                                                                                                HYPODENSE COMPARTMENT. OF PHOTORECEPTOR CELLS
                                                                                                                                                                                                                                                                           PHOTORECEPTOR CELL
                                                                                      REF.
                                                                                                                                                                                                                                                                                                                                               T BINDS 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --STPAGSSAN 1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -AHGTPSTPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein.";
                                                                                                                                                                                                                                                                                                                                                  E" TO
                                                                                                                                                                                                                                                                                                                                                                                              binds
                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1084
                                                                                                                                                                                                                                                                                                                                                   MOL
```

```
õ
                                                                                                                  δõ
                밁
                                                                                        В
                                                                                                                                                              В
                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                              Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 7.8%;
Best Local Similarity 27.5%;
Matches 216; Conservative 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Repeat.
DOMAIN
COMPLICT
COMPL
620 AKSPAEAKSPAEAKSPATVKSPGEAKSPSEAKSPA---EAKSPAEAKSPAEAKSPAEVKS
                                                                                                                          261
                                                                                        565
                                                                                                                                                510 TKSRVKEEAKSPGEAKSPGEAKSPA----EAKSPGEAKS-PGEAKSPGEAKSPAEPKSPA
                                                                                                                                                                                                                                                                                                                  404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                    201 AETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTTTKSAPTTPK
                                                                                                                                                                                                                          450 VIVEGOTEEIRVTEGVTEEEDKEAQGQEGEEAEEGEEKEEEELAAATSPPAEEAASPEKE
                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00038; filament; 1. PROSITE; PS00226; IF; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Intermediate filament;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001664; IF.
                                                                                                                                                                                                                                                                                                                                    86 DEAGSGLDNGDFKVT-----TPDTSTTQHNKVSTSPKITTAKPINPRPSLPPNSDTSKET 140
                                    ---PTTPKEPA----PTTPKEP----TPTTPKEPAPTTKEPAPTTPKEPA----PTAPKK
                                                                                                          EP-APTITKEPAPTTPKEPAPTTIKEP----APTITKSAPTTPKEPA----PTTPKKPA- 310
                                                                        EPKSPAEAKSPA--EPKSPA--TVKSPGEAKSPSEAKS-PAEAKSPAEAKSPAEAKSPAE
                                                                                                                                                                                                                                                          SLTVNKETTVETKETTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPK 200
                                                                                                                                                                                                                                                                                                  EECRIGFGPSPFSLTEGLPKIPSIST--HIKVKSEEMIKVVE------KSEKET 449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TO 716 AND IS SHORTER DUE TO FRAMESHIFTS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ₽
Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    116612 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coiled coil; Neurone; Phosphorylation;
                                                                                                                                                                                                                                                                                                                                                                                    81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOINED.
JOINED.
JOINED.
ALT_FRAME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LINKER 1.

COIL 1B.

LINKER 12.

COIL 2A.

COIL 2B.

K -> QA (IN REF. 2 AND 3).
A -> AR (IN REF. 2 AND 3).
A -> P (IN REF. 2 AND 3).
L -> G (IN REF. 2 AND 3).
L -> G (IN REF. 3).
G -> PREAKSP (IN REF. 3).
G -> A (IN REF. 3).
G -> M (IN REF. 3).
G -> A (IN REF. 3).
C -> M (IN REF. 3).
C -> A (IN REF. 3).
C -> A (IN REF. 3).
C -> A (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                        Score 533; DB 1; Length 1087;
Pred. No. 2.7e-15;
B1; Mismatches 330; Indels 158; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLU-RICH (ACIDIC).

50 x 6 AA TANDEM REPEATS OF K-S-P-A-E-A.
GLU/LYS-RICH.
COIL 1A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HEAD.
ROD.
TAIL.
                                      355
                                                                                                                                                                                      260
                                                                            619
                                                                                                                                                  564
                                                                                                                                                                                                                            509
                                                                                                                                                                                                                                                                                                                                                                            38;
```

CCCCCCCRTTRXX	RRI RA RA RRI	RN OCC	7.10	Db Qy	Db QY	Ор	ДУ	Db Qy	B 8	ob oy	ap da	Оу
SUBCELLULAR LOCATION. MEDILINE-91376547; PubMed-1896773; Prioli R.P., Mejla J.S., Aji T., Aikawa M., Pereira M.E.A.; "Trypanosoma cruzi: localization of neuraminidase on the surface of trypomastigotes."; Trop. Med. Parasitol. 42:146-150(1991). -! FUNCTION: DEVELOPMENTALLY REGULATED NEURAMINIDASE IMPLICATED IN PARASITE INVASION OF CELLS. -! CATALYTIC ACTIVITY: HYDROLYSIS OF 2,3-, 2,6- AND 2,8-GLYCOSIDIC LINKACES JOINING TERMINAL NON-REDUCING N- OR O-ACYLNEURAMINYL RESIDUES TO GALACTOSE, N-ACETYLHEXOSAMINE, OR N- OR O-ACYLATED NEURAMINYL RESIDUES IN OLIGOSACCHARIDES, GLYCOPROTEINS, GLYCOLIPIDS OR COLOMINIC ACID.	NCE FR N=SILV NE=912 ra M.E i R.P. i R.P. Trypan rial n tor, a p. Med	01-NOV-1991 (Rel. 20, Created) 01-NOV-1991 (Rel. 20, Last sequence update) 20-AUG-2001 (Rel. 40; Last annotation update) STALIDASE (EC 3.2.1.18) (NEURAMINIDASE) (NA) (MAJOR SURFACE ANTIGEN). TCNA. Trypanosoma cruzi. Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma. NCBI_TaxID=5693;	LT 9 _TRYCR TCNA_TI P23253	809 TITEK 813 	- ω - a	89 PKETAPTTPKEPAP1	629 KEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTT 688	571 APTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTP 628	49 RP	63 TKKPAPTAF 90 AKSPVKEDI		7 5

```
Qy
                                                                                                                                                                                   Вþ
                                                                                                                                                                                                         Qy
                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                             80
                                              δÃ
                                                                              D
                                                                                                   Qy
                                                                                                                           Вb
Ş
                      дb
                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation."; Genes Dev. 3:1639-1646(1989),
-i- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION. HAS A SPECIALIZED STRUCTURAL FUNCTION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nicotiana tabacum
Eukaryota; Viridip
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-1992 (Rel. :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EWBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-CV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HRGPNT3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Keller
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=90128263;
                                                                                                                                                                                                                                                                                                                                                            DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydroxylation.
                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                   141
                                                                                                                                                                                                                250
                                                                                                                                                                                     81
                                                                                                                                                                                                                                         24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLYCOSYLATED.
                                                                                                                                                                                                                                                                  AETTTK-GPALTTP--KEPTPTT----PKEPASTTPKE---PTPTTIKSAPTTPKEPAP
                                                                                                                                                                                     SHGHLPPSVGGPPPHRGHLPPSRGFNPPPSPVISPSHPPPSYGAPPPSHGPGHLPSHGQR
                                                                                                                                                                                                                                         AEATTQYGGYLPPPVTSQPPPSSIGLSPPSAPTTTPPSRGHVPSP---RHAPPRHAYPPP
                           RHLPPSPRRQPQPPTYSPPPPAYAQSPQPSPTYSPPPPTYSPPPPSPTYSPPPPAYSPSP
                                                     ---APTTPKEPSPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPT---TPKEPAPTTTK
                                                                                                        PAPTTPKEPAPTTPKEPAPTTTKEPSPTTPK---
                                                                                                                                   P-----PSPSHGHAPPSGGHTPPRGQHPPSHRRPSPPSRHGHPPPPTTYAQPPPTTIYS
                                                                                                                                                           PAPTTPKKPAPT - - - TPKEPAPTTPKEPTPTTPKEPAPTTK - - EPAPTTPKEPAPTAPKK
                                                                                                                                                                                                               TTTKSAPTTPKEP-----APTTTKEPAPT---TPKEPAPTTTKEPAPTTTKSAPTTPKE
                                                                               PSPQV--QPPPT--YSPPPPTHVQPTPSPPSRGHQPQPPTHRHAPPTHRHAPPTHQPSPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X13885; CAA32090.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B., Lamb C.J.;
                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ROOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             XANTHI;
                                                                                                                                                                                                                                                                                            8.1%;
ilarity 27.2%;
Conservative 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           wall;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S06733
    APKEPAPTTPKETAPTTPKKL-TPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21, Last annotation update)
R (CELL WALL HYDROXYPROLINE-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PubMed=2612909;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Common
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Leaf;
                                                                                                                                                                                                                                                                                                                                                                              620
73
151
242
235
242
620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                65406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tobacco)
                                                                                                                                                                                                                                                                                                                                                                WW;
                                                                                                                                                                                                                                                                                                64;
                                                                                                                                                                                                                                                                                                                                                                           2.
CONTAINS THE
                                                                                                                                                                                                                                                                                                                                                                                                                              H-A-P-P.
H-A-P-P.
2 X 7 AA
                                                                                                                                                                                                                                                                                                            Score 551; DB 1;
Pred. No. 3.1e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EXTENSIN
                                                                                                                                                                                                                                                                                                                                                            ONTAINS THE SER-PRO(4) REPEATS
X APPROXIMATE TANDEM REPEATS
641DD2278AB28524 CRC64;
                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                 A
                                                                                                                                                                                                                                                                                                                                                                                                                                   TANDEM
                                                                                                           -EPAPTTTKSAPTTTKEPAPTTTKS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Structural
                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEATS
        PEKLAPTTPEKPAPTTPEEL
                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLYCOPROTEIN)
                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              protein
                                                                                                                                                                                                                                                                                                                                                                                              REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                   OF T-H-R-H-A-P-P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    THE PE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PENTAPEPTIDE
                                                                                                                                                                                                                                                                                                    88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         THE
                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                              140
                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                           249
                                                              464
                                                                                                                 410
                                                                                                                                          193
                                                                                                                                                                     355
                                                                                                                                                                                                                                                80
                                    309
                                                                                      249
           514
                                                                                                                                                                                                                                                                                                       24;
```

```
RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
    δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      망
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    p19246; Q61959;
01-NOV-1990 (Rel. 16, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
20-AUG-2001 (Rel. 40, Last annotation update)
NEUROFILAMENT TRIPLET H PROTEIN (200 KDA NEUROFILAMENT
(NEUROFILAMENT HEAVY POLYPEPTIDE) (NF-H).
SUBMITTED (MAR-1994) to the EMBL/GenBank/DDBJ databases.
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
NP-H HAS AN IMPORTANT FUNCTION IN MATURE AXONS THAT IS NOT
SUBSERVED BY THE TWO SMALLER NF PROTEINS.
-I- PTM: THERE ARE A NUMBER OF REPEARS OF THE TRIPEPTIDE K-S-P, NFH IS
-I- PTM: THERE ARE A NUMBER OF THE SERINES IN THE FORMATION OF
THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF
INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE
OF AXONAL CALIBER.
-I- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING
OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE
LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND
COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION:
-I- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
-I- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 534
                                                                                                                                                                                                                                                                                                   MEDITINE-89089138; PubMed-3145094; Shneidman P.S., Carden M.J., Lees J.F., Lazzarini F The structure of the largest murine neurofilament revealed by cDNA and genomic sequences."; Brain Res. 464:217-231(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NFH_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      310
                                                                                                                                                                                                                                                                                                                                                                                                           Gene 68:307-314(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                           "Sequence and structure neurofilament subunit.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mushynski W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE=89121513; PubMed=3220257;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                572
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          515
                                                                                                                                                                                                                -i- FUNCTI
                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Julien J.-P., Cote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NEFH OR NFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             692
      ---
                                                                                                                                                       -
                                                                                                                                                                                                                                              Carden M.J.;
                                                                                                                                                                                                                                                          STRAIN-SWISS WEBSTER;
                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APTTPEEPTPTTPEEPAPTTPKAAAP----NTPKEPAPTTPKEPAPTTPKEPAPTTPKETA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TAPTTPKEPAPT -- TPKKPAPTTPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PPPSPIYSPPPDQVQPLPPTFSPPPPRRIHLPPPPHRQP----RPPTPTYGQPPSPPTFSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPPPPSSPPPPSFSPPPPTYEQSPPPPPAYSPPLPAPPTYSPPPPT---YSPPPPTYAQPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PSP------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTPKALENSPKEPGVPTTKTPAATKP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PPPRQIHSPPPPHWQPRTPTPTYGQPPSPPTFSAPPPRQIHSPPPPHRQPRPPTPTYGQP 598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---LPPTYSPPPPAYSPPPPTYSPPPPTYSPPPAYAQPPPPPTYSPPPAYSP
                                                                                                                                                                                                                                                                           FROM
                                                                                                                                                                                                                                                                             N.A.
                                                                                                                                                                                                                                                                                                                                                                                N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ۳J
:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -PTTYSPPSPPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chordata;
Rodentia;
                                                                                                                                                                                                                                                               TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Beaudet
                                                                                                                                                                                                                                                                                                                                                                                                                                            of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Craniata; Veri
Sciurognathi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ľ.,
                                                                                                                                                                                                                                                                                                                                                                                                                                              mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sidky
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1087
                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Vertebrata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A
                                                                                                                                                                                                                                                                                                                                                                                                                                              coding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Muridae; Murinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Flavell
                                                                                                                                                                                                                                                                                                                                                                                                                                               for
                                                                                                                                                                                                                                                                                                                                                         R.A.;
                                                                                                                                                                                                                                                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                 the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROTEIN)
                                                                                     THE FUNCTIONING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Grosveld
                                                                                                                    MAINTENANCE
                                                                                                                                                                                                                                                                                                                                            (NF-H)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ή
.
```

```
THE THE TERM DRAWER DRA
                                                                                                                                                                                                                                                                               SMARR; SM00137; MAM; 2

SMART; SM00214; VWC; 1

SMART; SM00011; VWC_def; 3

SMART; SM00216; VWD; 3

PROSITE; PS01186; EGF2; 3

PROSITE; PS00740; MAM_1; 1

PROSITE; PS50060; MAM_2; 4
          DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIGNALING.

SIGNALING.

SIGNALING.

SIGNALING.

SIGNALING.

SUBGLILULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).

APICAL REGION OF THE SPERM HEAD (BY SIMILARITY) IN THE STATE SPERM ADDIATE SPERM ADHESION TO THE DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE ZONA PELLUCIDA.

DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS, THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.

SPERMATOZOA OR PROMOTING ADHESION TO THE COVALENT OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).

SIMILARITY: CONTAINS 3.5 MAM DOMAINS.

SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF00629; MAM; 3. Pfam; PF001826; TIL; 4. Pfam; PF003845; TILa; 4. Pfam; PF00094; vwd; 4. PRINTS; PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                   DOMAIN
                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AF053356; AAC78790.1; -. EMBL; U83191; AAC51208.1; -. MIM; 602372; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gao Z., Harumi T., Garbers D.L.;
"Chromosome localization of the mouse zonadhesin zonadhesin gene (ZAN).";
Genomics 41:119-122(1997).
-i- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Large-scale sequencing of two regions in numerical controlled analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci reveals 17 genes.";
Genome Res. 8:1060-1073(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-97271566; PubMed-9126492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 2338-2700 FROM N.A. TISSUE-Testis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=99018118; PubMed=9799793; Glockner G., Scherer S., Schattev Tsui L.C., Rosenthal A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ); IPR000561; EGF-like.
); IPR000998; MAM.
); IPR002996; P_rich_ext.
); IPR002919; TIL.
); IPR003328; TILA.
); IPR001007; VWFC.
); IPR001846; Vwd.
          953
1066
1455
                                                                                                           1
112
161
322
483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Rosenthal A.;
                                                                                                                                                                                                                                                            Transmembrane;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Scherer S., Schattevoy
        1065
1454
1861
                                                                                                           1
109
136
326
446
951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P_rich_extensn.
MAM 2 (PARTIAL).
MAM 3.
MAM 4.
66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
(MUCIN-LIKE DOMAIN).
VWED 1 (PARTIAL).
VWED 2.
VWFD 3.
                                                                                                                                                                                                                                                               Cell
                                                                                                                                                                                                                                                         adhesion;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R., Boright A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MANNER TO THE ZONA RECOGNITION AND/OR
                                                                                                                                                                                                                                                      Repeat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Weber J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         the
```

RESULT 7
EXTN_TOBAC
ID EXTN_T
AC P13983
DT 01-JAN
DT 01-JAN

STANDARD;

620

EXTN_TOBAC P13983; 01-JAN-1990 01-JAN-1990

(Rel. 13, Created) (Rel. 13, Last sequence

```
В
              δÃ
                                                   Q
                                                                        밁
                                                                                       Š
                                                                                                           밁
                                                                                                                       .0
                                                                                                                                              В
                                                                                                                                                              QУ
                                                                                                                                                                                  В
                                                                                                                                                                                                   Q
                                                                                                                                                                                                                      В
                                                                                                                                                                                                                                     δÃ
                                                                                                                                                                                                                                                         B
                                                                                                                                                                                                                                                                        ρy
                                                                                                                                                                                                                                                                                            В
                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                Ş
                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                      ρ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CONFLICT
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
DOMAIN
DOMAIN
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
967
                 694
                                    912
                                                     636
                                                                        856
                                                                                        582
                                                                                                           815
                                                                                                                            522
                                                                                                                                              758
                                                                                                                                                                472
                                                                                                                                                                                  702
                                                                                                                                                                                                    417
                                                                                                                                                                                                                      648
                                                                                                                                                                                                                                       357
                                                                                                                                                                                                                                                         597
                                                                                                                                                                                                                                                                           297
                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                 488
                                                                                                                                                                                                                                                                                                                                                                                      123
                                                                                                                                                                                                                                                                                                                                                 183
                                                                                                                                                                                                                                                                                                                                                                                              Ma
Local
199;
PASCKSPRPS
                                                               PTTPKEPAPT 703
                            PPHPSPTA-TGLAALVMSPHAPSTPMTSV---ILGTTTTSRSSTGMSCP-PNARYESCAC
                                             PKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKET--A
                                                                                                    KPTISPEKPTISTEK------PTIPTE-KPTIPTE-----
                                                                                                                EKPTISTEKPTIPTEKPTISPEKLTIPT-EKLTIPTEKPTIPIEETTISTEKL--TIPTE
                                                                                                                                                         KEPAPTTPKETAPT----TPKKLTPTTPEKLA----PTTPEKPAPTTPEELAPTTPEE
                                                                                                                                                                          EKPTISTEE--PTTPTEETTISTEKPSIPMEKPTLPTEETTTSVEETTISTEKLTIPM--
                                                                                                                                                                                      EPSPTTTKEPAPTTPKEPAPTTPKKPA-----PTTPKEPAPTTPKEPAPTTTKKPAPTAP
                                                                                                                                                                                                                -PTTPTE-KPTIPTEKPTISTEKPTIPTEK-PTISPEKLTIPTEKLTIPT---EKPTIPT
                                                                                                                                                                                                                         APTTPKEPAPTTPKEPAPTTTKKEPSPTTPKEPAPTTTKSAPTTTKSAPTTPK 416
                                                                                                                                                                                                                                                  --EKPSIPTEK---PTISMEETIISTEKPTICPEKPTIPTEK----PTIPTEKSTISPEK-
                                                                                                                                                                                                                                                                   TPKEPAPTTPKKPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPTAPKKP 356
                                                                                                                                                                                                                                                                                     TEKPTIPSE-KPTIPSEKPTISTEKPTVPTEE--PTTPTEETTTYMEEPVIPTEKPSIPT
                                                                                                                                                                                                                                                                                                  KSAPTTPKEPAPTTTKSAPTTPKEPAPTTTKEPAPTTTKEPA-PTTTKSAPT
                                                                                                                                                                                                                                                                                                                          SAKDLAPTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTPKEPT-----PTTI
                                                                                                                                                                                                                                                                                                                                                            PVKVLPELPPVSS----TGPSETTGLTENPTISTK-----KPTVSIEKP
                                                                                                                                                                                                                                                                                                                                                                        PINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKT 182
                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74
403
1023
1099
1618
1737
1832
1878
                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
976
                                                                                                                                                                                                                                                                                                                                                                                                                                               2700
1; 293013 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74
403
1023
1099
1618
1737
1832
1878
2136
2505
                                                                                                                                                                                                                                                                                                                                                                                                            8.2%;
                                                                                                                                                                                                                                                                                                                                                                                                    64;
                                                                                                                                                                                                                                                                                                                                                                                                 Score 555.5; DB
Pred. No. 7e-16;
64; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VWFD 4.
VWAD 5.
EGF-LIKE.
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                               80E60CC0B12277B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
(GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RAGPGP
                                                                                                                                                                                                                                                                                                                                                                                                                     DB
                                                                                                                                                                                                                                                                                                                                                                                                   240;
                                                                                                                                                                                                                                                                                                                                                                                                                   ۲.
                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                REF.
                                                                                                     -ETTISTEKLTIPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                  2700;
                                                                                                                                                                                                                                                                                                                                                                                                 107;
                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                               693
                                                                911
                                                                                                                                                         521
                                                                                   635
                                                                                                    855
                                                                                                                     581
                                                                                                                                       814
                                                                                                                                                                                              471
                                                                                                                                                                          757
                                                                                                                                                                                                                701
                                                                                                                                                                                                                                                   647
                                                                                                                                                                                                                                                                                                        296
                                                                                                                                                                                                                                                                                                                          539
                                                                                                                                                                                                                                                                                                                                            237
                                                                                                                                                                                                                                                                                                                                                              487
                                                                                                                                                                                                                                                                                                                                                                                                32;
```

```
l_MXIL., FPI_MYTED
, FPI_MYTED
, Q25460;
Q25460;
T 30-MAY-2000 (Rel. 39, Created)
T 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MG-2001 (Rel. 40, Last annotation update)
DT 20-MG-2001 (Rel. 40, Last annotation update)
ADHESIVE PLAQUE MATRIX PROTEIN (POLYPHENOLIC AD ADHESIVE PLAQUE P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                    δÃ
                                                                                                                                  Дb
                                                                                                                                                                                              ΛÕ
Вþ
                                                                                                                                                                                                                                                                                 Query Match
Best Local S
Matches 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MYTED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.

MEDLINE-91025029; PubMed=1367451;

Filpula D.R., Lee S.M., Link R.P., Str

"Structural and functional repetition
protein.";

Biotechnol. Prog. 6:171-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mytilus edulis (Blue mussel).
Eukaryota; Metazoa; Mollusca;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Evidence for a repeating 3,4-dihydroxyphenylalanine- and hydroxyproline-containing decapeptide in the adhesive protein mussel, Mytilus edulis L.";
J. Biol. Chem. 258:2911-2915(1983).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mytiloidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=6550;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way use by non-profit institutions as long as its content is in no way use by non-profit institutions as long as its content is in no way use by non-profit institutions as long as its content is in no way as the content is in the content in the content in the content is in the content in the content in the content is in the content in the content in the content is in the content in t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Waite J.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PARTIAL SEQUENCE, MEDLINE=83135732;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; X54422; CAA38294.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                       153
                                                                                                                                                                                                                                                                             Local Similarity
mes 272; Conserv
   60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BIOI. Chem. 258:2911-2915(1983).

FUNCTION: PROVIDES ADHESTVENESS TO THE MUSSEL'S FOOT. MUSSELS PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S PRODUCE ONE OF THE STRONGEST WATER ADHESTVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A FIBROUS COLLAGENOUS CORE COATED WITH ADHESTVE PROTEINS.

SUBCELLULAR LOCATION: SECRETED

TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

TOMAIN: ALMOST EXCLUSIVELY COMPOSED OF A DECAPEPTIDE.

DOMAIN: THE DECAPEPTIDE A-K-P-S-Y-P-T-Y-K IS POST-TRANSLATIONALLY

MODIFIED AS FOLLOWING: THE SIXTH AND SEVENTH RESIDUES ARE

MYDROXYLATED AND THE PENULTIMATE IS A 3,4-DIHYDROXYPHENYLALANINE
                                                                                                                                              Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOPA) DERIVED FROM TYROSINE.
                                                                   KETTTTNKQTSTDG--KEKTT---SAKETQSIEKTSAKDLAPTSKVLAKPTPKAETTTKG
                                                                                                                                          TKHEPYYKPYKTSYSAPYKPPTYQPLKKKVDYRPTKSYPPTYG-SKTNYLPLAKKLSSYK
                                                                                                                                                                                                                 TQHNKVSTSPKITTAKPINP---
PIKTTYNAKTNYPPVYKPKMTYPPTYKPKPSYPPTYKSKPTYKPKITYPPTYKAKPSY--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            o; IPR002964; Adhesive_plaq.
b; IPR002965; P_rich_extensn.
PR01216; ADHESIVEI.
PR01217; PRICHEXTENSN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hydroxylation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mytilidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    875
                                                                                                                                                                                                                                                                                            Conservative 106; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ΑĀ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PubMed=6298211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AND POST-TRANSLATIONAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100412 MW;
                                                                                                                                                                                                                                                                                                                             8.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mytilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bivalvia;
                                                                                                                                                                                                                                                                                                Score 556.5; DB 1;
Pred. No. 2.5e-16;
6; Mismatches 375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TANDEM REPEATS OF Y-K-[PA]-K-[LP]-[ST]-Y-P--[ST].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6EA85312748CAACE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Strausberg S.L., Strausberg R.L.;
on in a marine mussel adhesive
                                                                                                                                                                                                                        RP--SLPPNSDTSKETSLTVNKETTVET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pteriomorphia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MODIFICATIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADHESIVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CRC64;
                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mytiloida;
                                                                                                                                                                                                                                                                                                           195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ō,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (FOOT
                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           the
                                                                                                                                                                    59
                                                                                             207
                                                                                                                                                                                                                                                                                                              56;
```

```
Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          밁
                                                                                                                                                                                                                                                                                                QУ
                                                                                                                                                                                                                                                                                                                               Вb
                                                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                      Вþ
                                                                                                                                                                                                                                                                                                                                                                                                             γ
                                                                                                                                                                                                                                                                                                                                                                                                                                            Б
                                                                                                                                                                                                                                             Qy
                                                                                                                                                                                                                                                                          Вþ
                                                                                                                                                                                      Qy
                                                                                                                                                                                                                 DЪ
B
                       Qy
                                                     Dр
                                                                            QY
                                                                                                          В
                                                                                                                              Qγ
                                                                                                                                                               멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               259
                                                                                                                                                                                                                                                                                                                                                                                       458
                                                                                                                                                                                                                                                                                                                                                                                                                                             410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          306
                                                                                                                                                                                                                                                                                                                                                                                                                   521
                                                                                                                                                                 668
                                                                                                                                                                                                                      612
                                                                                                                                                                                                                                                 683
                                                                                                                                                                                                                                                                            552
                                                                                                                                                                                                                                                                                                      631
                                                                                                                                                                                                                                                                                                                                  507
                                                                                841
                                                                                                          727
                                                                                                                                      781
 827
                           868
                                                      778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PALTTPKEPTPTTPKE----PASTTPKEPTPTTIKSAPTTP----KEPAPTTTKSAPTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKPTYKAKPSYPPTYKA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PKKPA--PTTPKEPA--PTTPKEPT--PTTPKEP---APTTKEPAPTTPKEPA--PTAPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P----PTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYPSTYKAKPSYPPTYKAKPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PSSYKPKKTYPPTYKPKLTYPPTYKPKPSYPPTYKPKPSYPPSYKTKKTYPSSYKAKPSY
                                                                                                                                                                                                                                                                                                                                                                                                         EPTPTTPEEPAPTTPKAAAAPNTPKEPAPTTPKEPAPTTPKE----PAPTTPKETAPTT--
                                                                                                                                                                                                                                                                                                                                                                                                                                           PPTYKAKPTYKAKP-----TYPST-YKAKPSYP----PSYKAKPSYPPTYKAKRTYKA 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPTYKAKPTYKAKPTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KPA--PTTPKEPA--PTTPKEPA--PTTTKEPA-PTTTKSAPTTTKEPA-
||: || :|: || :|: || :| || :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: || :|: ||
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PKEPAPTTTKEPA -- PTTPKEPA -- -- --
                                                                                                                                                                                                                                                                                                   PAPTTPKEPAPTTPKGT---APTTLKEPA--PTTPKKPA--PKELAPTTTKG-PTSTTSD
                                                                                                                                                                                                                                                                                                                                 PKLTYKPTYK-PKPSYRPKYKPKTTYPDTYK------PKISYPPTYKAKPSY
                                                                                                                                                                                                                                                                                                                                                          PKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAPTTPKGTAPTT----PKE
                                                                                                                                                                                                                                                                                                                                                                                      KPT-----YPSTYKA----KPSYPASYKAKPSYPPTYKSKSSYPSSYKPKKTYPPTYK 506
                                                                               TLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAPKK---P
                                                                                                                                    AKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKIT
                                                                                                                                                                                                                     KAKPSYP----PTYKAKPSYPPTYKAKPTYPSTYKAKPSYPPTYKPKISYPPTYKAKPSY
                                                                                                                                                                                                                                              KPAPTTPKETAPTTPKEPA--PTTPKKPA-PTT----PETPPPTTSEVSTPTTTKE----
                                                                                                                                                                                                                                                                          PATYKAKPSYPPTYKAKPSYPPTYKAKPSYPPTYKAKPSYKAKPTYPSTYKAKPSYPPTY
                                                                                                                                                                PPT-YKAKPSYPPTYKAKPTYKAKPTNPSTYKAKPSYPPTYKAKPSYPPTYKAKPSYPPT
                                                                                                                                                                                        PTTIHKSPDESTPELSAEPTPKALENSPK----EPGVPTT--KTPA-----ATKPEMTTT
                                                     TYKSKSIYPSSYKPKKTYPPT---YKPKLTYPPTYK--
                                                                                                           YKAKPTYKAKPTYPSTYKAKP--TYKAKPTYPPTYKAK---
 STYKLKPSYPPTYKSKTSYPPTYNKKISYPSSYKAKTSYPPAYKPTNR
                          TSTKKPKTMPRVRKPKTTPTP---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --PTTTKEPA-PTTTKSAPTTPK--EPAPTT
                              RKMTSTMPELNPTSR
                                                            PKPSYPPSYKPKITYP
                                                                                                                  -PSYPPTYKPKPSYPP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          293
                                                                                                                                                                                                                                                                                                                                           551
                                                                                                                                                                                                                                                                                                                                                                    630
                                                                                                                                                                                                                                                                                                                                                                                                                          574
                                                                                                                                                                                                                                                                                   611
                                                                                                                                              840
                                                                                                                                                                                                  780
                                                                                                                                                                                                                              667
                                                                                                                                                                                                                                                          731
                                                                                                                                                                                                                                                                                                               682
                                                               826
                                                                                           897
                                                                                                                    777
                                                                                                                                                                         726
```

```
RESULT (
 RP RN
                                                                  ZAN_HUMAN STANDARI
Q9Y493; 000218;
20-AUG-2001 (Rel. 40, 0
20-AUG-2001 (Rel. 40, 1
20-AUG-2001 (Rel. 40, 1
20-AUG-2001 (Rel. 40, 1
20NADHESIN (FRAGMENT).
                            Homo sapiens (Human)
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                           ZAN
SEQUENCE OF 1-2379
                  NCBI_TaxID=9606;
                                                                                                                                         6
                                                                                                                    STANDARD;
   FROM
                                                                               Created)
Last sequence up
   N.A.
                              Craniata; Vertebrata; Catarrhini; Hominidae;
                                                                                                                        PRT;
                                                                                                                        2700
                                                                                          update)
                                                                                update)
                                                                                                                        B
```

Euteleostomi;

```
Interpro; 115/24;

R Interpro; 1PROMON82; SEA; 1.

R SMART; SMOO200; SEA; 1.

R SMART; SMOO200; SEA; 1.

R PROSITE; PS50024; SEA; 1.

R Repeat; Signal; Cytoskeleton; Actin-binding; Transmembrane; Gycoprotein; Signal; Cytoskeleton; Actin-binding; Transmembrane; ET SIGNAL

R Repeat; Alternative splicing.

POTENTIAL.

P
  Query Match 9.6%; Score 651; DB 1; Best Local Similarity 27.8%; Pred. No. 5.1e-20; Matches 295; Conservative 72; Mismatches 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARSPLIC
CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                              VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TUMORS, SUCH AS BREAST CANCER.

PTM: HIGHLY GLYCOSYLATED (N-AND O-LINKED CARBOHYDRATES AND SIALIC ACID).

ACID).

POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC. IT VARIES FROM 21 TO 125 IN THE NORTHERN EUROPEAN POPULATION. THE MOST FREQUENT ALLELES CONTAINS 41 AND 85 REPEATS.

SIMILARITY: CONTAINS 1 SEA DOMAIN.
                                                                                                        134
154
1021
1251
1255
                                                                                                                                                                                                                       1088
                                                                                                                                                                                                                                                              1077
                                                                                                          AA,
                                                                                                                       1255
134
154
154
1021
1251
                                                                                                                                                                                                                                                            1087
                                                                                                        122072
                                                                                          SEA.

N-LINKED (GLCNAC...) (POTENTIAL).

T-STATTAPKPAT (IN ISOFORM B).

MISSING (IN ISOFORM D).

MISSING (IN ISOFORM D).

MISSING (IN ISOFORM D).

MISSING (IN POLYMORPHIC EPITHELIAL ISOFORM).

FLQIVKQGGFL -> VSIGLSFPMLP (IN SECRETED ISOFORM).

T-> A (IN REF. 11).

P-> Q (IN REF. 9).

S-> T (IN REF. 3).

A-> T (IN REF. 3).

P-> Q (IN REF. 3).

SEZBDFC4DE7D9A82 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALT_INIT.
ALT_INIT.
ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALT_SEQ.
ALT_SEQ.
ALT_SEQ.
      460;
                                         Length 1255;
  Indels 234;
  Gaps
51;
```

	1006 SDETNICNGKPVDGLTTLRNGTLVAFRGHYFWMLSPFSPPS 1046 : :	Оy
955		ДЬ
1005	AEGETPHMLLRPHVFMPEVTPDMDYLPRVPNOGITINDML	Qy
946 909	00/ PURTKAPKRETSTKREKTMERFKRETSTERFKMTSTMEELNETSRIABAMLQTTTREN (dd Gy
867	PGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPA	₽ ₽
886		Qγ
807	PPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPA	DЪ
، س	TTESKITATTTOVTSTT	Qy
753		Db
777	TKTPAATKPEM	Qy
700	PAHGVTSAPDTRPAPGSTAPPAHGVTS	DЬ
734	KKDAPTTDETDDTTCEVC	Оу
642	RPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAP-PAHGVTSAP	Db
680	ELAPTTTKGPTSTT	Qy
583	: STAPPAHGVTSAPD	ф
634	-TTPKGTAPTTPKEPAPT	VΩ
528		Db
586	APNTPKEPAPTTPKEPAPTTPKEPAPTTP	Qy
472	 RPAPGSTAP	Db
5 . 2 .	TTPEKPAPTTPEELAPTTPEEPTPTTPEEPA	Qy
416	PPAHGV	B B
474	-EPAPTTTKKPAPTAP	Qy
35 6	: PGSTAP	В
	TTKEPAPTTTKSAPTTPKEPSPTTTK	Qy
	 GSTAPPAHGV	рь
37	TTPK	Qy
23	PAHGV	Db
ω	KKPAPTTPKEPAPTTPKEPTPTTP	Qy
17	126 PAPGSTAPPAHGVT-SAPDTRPAPGSTAPPAHGVTSAPDTRPAPGSTAPPAHGVT-	Db
29	SAPTTPKEPAPTTKEP	Qy
		Дb
N (GPALTTPKEPTPTTPKEPASTTPKEPTPTTI-	Qy
on 1	16 TVLTVVTGSGHASSTPGGEKETSATQRSSVPSSTEKNAVSMTSSVLSSHS	Дb
198	140 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKD-LAPTSKVLAKPT	Qy

```
δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
   RACE LINE REPORT TO THE REPORT OF THE REPORT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Вb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MUC1_HUMAN STANDARD; PRT; 1255 AA.

P15941; P15942; P13931; P17626; Q14128; Q16442; Q16437; Q9Y4J2;

P15941; P15942; P13931; P17626; Q14128; Q16442; Q16437; Q9Y4J2;

P15941; P15942; P13931; P17626; Q14128; Q16442; Q16437; Q9Y4J2;

P15941; P15942; P13931; P17626; Q16442; Q16442; Q164437; Q9Y4J2;

P15941; P15942; P15942; Q16442; Q16442; Q164437; Q9Y4J2;

P15941; P15942; P15942; Q16442; Q164437; Q9Y4J2;

P15941; P15942; Q16442; Q16442; Q164437; Q9Y4J2;

P15941; P15942; Q16442; Q16442; Q16442; Q164437; Q9Y4J2;

P15941; P15942; Q16442; Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1022
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
MCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-90202794; PubMed-2318825;
Ligtenberg M.J.L., Vos H.L., Gennissen A.M.C., Hilkens J.;
"Episialin, a carcinoma-associated mucin, is generated by a
polymorphic gene encoding splice variants with alternative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIINE=90368716; PubMed=2394722;
Lan M.S., Batra S.K., Ql W.-N., Metzgar R.S.,
"Cloning and sequencing of a human pancreatic
J. Biol. Chem. 265:15294-15299(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  978
                                                                                                                                                                                                                                                                                                  Gendler S.J., Lancaster C.A., Taylor-Papadimitriou J., Duhi; peat N., Burchell J., Pemberton L., Lalani E. N., Wilson D., "Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin."; J. Biol. Chem. 265:15286-15293(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE-Pancreas;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
"Structure and expression of the human gene: an expressed VNTR unit."; Biochem. Biophys. Res. Commun. 173:1019
                                                                                                                 MEDLINE=91097524; PubMed=2268309;
Lancaster C.A., Peat N., Duhig T., Wilson
Taylor-Papadimitriou J., Gendler S.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. TISSUE=Breast carcinoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      termini.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         polymorphic gene encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=90368715; PubMed=1697589;
                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Biol. Chem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    YPGSQTETSVSSTTETTIVPTKTTTSVTTPSTTTITTTVCSTGTNSAGETTSGCSPKTVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TTVPTTTTTSVTTSSTTTITTTVCSTGTNSAGETTSGCSPKTITTTVPCSTSPSETA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AT--TTQVTSTTTQDTTPFKITTLKTTTLAPKVTT---TKKTITTT-EIMNKPEETAKPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TLAPSAPVTPATN----AVPTTITT--TECSAATNAAGET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STGTGEYTTEATTLYTTAVTTTVVTTESSTGTNSA-GKTTTGYTTKSVPTT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HVFMPEVTPDMDYLPRVPNQGIIINPMLSDETNICNGKPVDGLTTLRNGTLVAFRGHYFW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P--SVTTVTNFTPTT-ITTTVCSTGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DRATNSKATTPKPQKP - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MLSPFSPPSPARRITEVWGIPSPIDTVFTRC----NCEGKT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -SESTTTSPTTPVTTVVSTTVVTTEYSTSTKPGGEITTTFVTKNIPTTYLTTIAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       265:5573-5578(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             splice
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Craniata; Vertebrata; Catarrhini; Hominidae
                   173:1019-1029(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -TKAPKKPTSTKKPKTMPRVRKPKTTPT
                                                                                          polymorphic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hominidae;
                                                                                                                                                                    .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -NSAGETTSGCSPKTVTTTVPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hollingsworth M.A.;
tumor mucin cDNA.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1074
                                                                                              epithelial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Euteleostom1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Duhig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TVY-
                                                                                                         mucin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          977
```

Hareuveni M., Tsarfaty I., Zaretsky J., Kotkes P., Horev J., Zrihan S., Weiss M., Green S., Lathe R., Keydar I., Wreschner D.H.; A transcribed gene, containing a variable number of tandem repeats, codes for a human epithelial tumor antigen. cDNA cloning, expression of the transfected gene and over expression in breast cancer Horov J., Zaretsky J., Kotkes P., Weiss M., Lathe R., Dion A., Keydar I.;
"Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms.";
Eur. J. Biochem. 189:463-473(1990). MEDLINE-90276413; PubMed-2351132; Wreschener D.H., Hareuveni M., Tsarfaty Horov J., Zaretsky J., Kotkes P., Weiss [5] SEQUENCE FROM Tsarfaty I., Hareuveni M., Horev J., Z Jeltsch J.M., Garnier J.M., Lathe R., "Isolation and characterization of an coding for a breast cancer-associated SEQUENCE FROM N.A.
TISSUE=Breast carcinoma;
MEDLINE=90276414; PubMed=2112460; TISSUE=Breast carcinoma, IS ALSO PRODUCED.
-!- ALTERNATIVE PRODUC
SPLICING.
-!- TISSUE SPECIFICITY "A highly immunogenic region of a human polynexpressed by carcinomas is made up of tandem J. Biol. Chem. 263:12820-12823(1988). MEDLINE=91033045; PubMed=1688329; SEQUENCE FROM N.A. tissue. Abe M., Siddiqui J., Kufe D.; "Sequence analysis of the 5'; SEQUENCE OF 1-169 FROM N.A. MEDLINE=90088473; PubMed=2597151; Gene TISSUE=Thyroid; MEDLINE=96183746; Biochem. Biophys. Res. Burchell J. MEDLINE=88330762; PARTIAL Weiss M., Baruch A., Keydar I., Wreschner D.H.; "Preoperative diagnosis of thyroid papillary carcinoma by transcriptase polymerase chain reaction of the MUC1 gene." Int. J. Cancer 66:55-59(1996). SEQUENCE OF 1-109 carcinoma-associated antigen gene [9] Gendler S.J., Buluwela L., Liu Q., Luqmani Y.A., Gomm J.J., Submitted (OCT-1992) to the EMBL/GenBank/DDBJ-i- FUNCTION: DIRECT OR INDIRECT INTERACTION W MEDLINE-96181716; PubMed-8604237; Yu C.J., Yang P.C., Shew J.Y., Ho Yu C.E. L.N., Luh K.T., Wu C.W.; "Mucin mRNA expression in lung ad SEQUENCE OF 1-89 FROM N.A. [10] TISSUE=Breast carcinoma, SEQUENCE OF Oncology tissues." J. Biochem. 189:475-486(1990). SUBCELLULAR LOCATION: 93:313-318(1990) CYTOSKELETON. SEQUENCE FROM N.A. =88330762; PubMed=3417635; 53:118-126(1996). SPECIFICITY: 1-46 FROM N.A. Taylor-Papadimitriou PRODUCTS: VARIOUS PubMed=8608966; FROM N.A. Commun. ABERRANTLY TYPE region 165:644-649(1989) Hong adenocarcinoma MEMBRANE J., VARIANTS ARE PRODUCED of EXPRESSED Zaretsky J., , Keydar I., T.M., Keydar I., Wreschner D.H
expressed hypervariable
antigen."; Duhig ж I the human polymorphic Yang PROTEIN. Smorodinsky Τ., repeats J databases.
WITH ACTIN WITH H s.c., , Weiss M., HUMAN Rothbard DF3 Þ lines epithelial s.": Lee breast SECRETED EPITHELIAL BY ALTERNATIVE reverse U . H.; gene FORM mucin

```
RESULT 3

ANYH, YAN 3

ANYH, YAN 3

ANYH, YAN 3

AC P08640

D1 01-FEE

D1 20-AUG

D2 01-FEE

D7 20-AUG

D7 01-FEE

D7 20-AUG

D8 GLUCOS

GN STA1 C

OS Saccha

OC Suckary

OC Saccha

OC Saccha

OC Saccha

OC Saccha

OC Saccha

OC Saccha

OC STA1 C

RA Gentle

RA Church

RA Gentle

RA Louis

RA Rajan

RA Harrel

RA Houis

RA Rajan

RA Houis

RA Walsh

RI Submit

RA Walsh

RI SUBMit

RA Walsh

RI STA1 C

RA WALSh

RI FEBS I

CC -1- C

CC -1- C

CC This |

CC Th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MYH_YEAST STANDARW;

D AMYH_YEAST STANDARW;

CC P08640; P08068;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-FBB-1995 (Rel. 31, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DT 20-AUG-2001 (Rel. 41, Last annotation update)

DE GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-B-GLUCAN GLUCOHYDROLASE).

OS GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE).

STA1 OR STA2 OR MAL5 OR YIR019C.

OS Saccharomyces cerevislae (Baker's yeast).

OC Eukaryota; Fung1; Ascomycota; Saccharomycete; cancharomycetales; Saccharomycetaceae; Saccharomycete.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                 MEDLINE-89031230; pubMed-3141213;
Pardo J.M., Ianez E., Zalacain M., Claros M.G.,
"Similar short elements in the 5' regions of the
from Saccharomyces cerevisiae.";
FEBS Lett. 239:179-184(1988)
-:- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL 1
GLUCOSE RESIDUES SUCCESSIVELY FROM NON-REDUC
WITH RELEASE OF BETA-D-GLUCOSE.
-:- SIMILARITY: TO S.POMBE SPBC215.13.
-:- SIMILARITY: SOME, TO S.POMBE SPCC285.13C.
                                                                  This SWISS-PROT entry is copyright. It is produced through a content the Swiss Institute of Bioinformatics and the EMBL the European Bioinformatics Institute. There are no restrict use by non-profit institutions as its content is modified and this statement is not removed. Usage by and fo entitles requires a license agreement (See http://www.isb-sib.or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-SSBBC / AB972;

Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D. Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., F. Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D. Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V. Walsh S.V., Whitehead S.;

Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1-31 FROM N.A. STRAIN-SPX101-1C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 1-242 AND 762-1331 FROM N.A. MEDLINE-87194600; Pubmed-3106330; Yamashita I., Nakamura M., Fukul S.; "Gene fusion is a possible mechanism undersal".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 1-242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1601 YIKRSESV----ALINRALERGPLNGAPKLFPDVNESYWAF-GDIMD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1454 YLRGYPDGSFRPERNITRAEAAVIF-----AKLL----GADESYGAQSASPYSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteriol. 169:2142-2149(1987).
Z38061; CAA86176.1; -. M16164; AAA35014.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PARRITEVWGIPSPIDTVFTRCNCEGKTFFFKD--SQYWRFTNDIKD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TIDISNPKFDD----CVGHWAQEFIEKLTSLGYISGYPDGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---IINPMLSDETNICNG----KPVDGLTTL-----RNGTLVAFRGHYFWMLSPFSPPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LADTHWAAWAIKFATSQGLFKGYPDGTFKPDQNITRAEFATVVLHFLTKVKGQEIMSKLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LLRPH------VFMPE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----TSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHM-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    underlying
                                                                                             moved. Usage by and for commercial (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                        TERMINAL 1,4-LINKED NON-REDUCING ENDS C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----VTPDMDYLPRVPNQGI-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     M.G., Jimenez A of the STA2 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 evolution
                                                                                                                                                                                 restrictions
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ۷ .
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fraser
                                                                                                                                                                                                                                                                                                                                                                                        OF THE CH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jones
                                                                                                                                                                                                                                                                                                                                                                                     THE
                                                                                                                                                                                                                                        collaboration
                                                                                                                                                                                                            outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of
                                                                                                                                                                                                                                                                                                                                                                                     CHAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X ?
                                                                                                                                                      0 On
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           973
```

```
Ş
                                  В
                                                                                    DЬ
                                                                                                           δõ
                                                                                                                                       В
                                                                                                                                                                                        DЬ
                                                                                                                                                             Q
                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                            B
                                                                                                                                                                                                                                                                                               В
                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                 망
                                                                                                                                                                                                                                                                                                                                                                         20
                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DR FTT FTT SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 313; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M16165; AAA35015.1;
EMBL; X13857; CAA32069.1;
PIR; B26877; B26877.
PIR; A26877; A26877.
PIR; A26877; A26877.
PIR; S48478; S48478.
SGD; S0001458; MUCLI.
Hydrolase; Glycosidase; Po:
Signal; Multigene family.
SIGNAL 1 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
SEQUENCE
         761
                                  849
                                                                                                              651
                                                                                    789
                                                                                                                                                                595
                                                                                                                                                                                         678
                                                                                                                                                                                                                  536
                                                                                                                                                                                                                                            623
                                                                                                                                                                                                                                                                    482
                                                                                                                                                                                                                                                                                               567
                                                                                                                                                                                                                                                                                                                        427
                                                                                                                                                                                                                                                                                                                                                   508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                    452
                                                                                                                                                                                                                                                                                                                                                                                                                               318
                                                                                                                                                                                                                                                                                                                                                                                                                                                        403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44 DCDAQCKKYDK---CCPDYESFCAEVK-DNKKNRTKKKPTPKPPVVDE--AGSGLDNGDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98
EPGVPTTKTPAATKPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTESKIT
                        SSSAPVSSSTTESSVAPVPTPSSSSNITSSAPSSI---PFSSTTESFSTGT-TVTPSSSK
                                                KKPAPTTPETPPPTTSEVSTPT-----TTKEPTTIHKSPDESTPELSAEPTPKALENSPK
                                                                          SSSAPVPTPSSSTTESSVAPVPTPSSSSNITSSAPSSTPFSSSTESSSVPVPTPSSSTTE
                                                                                                   TLKEPAPTTPKKPAPKELAPTTTKGPTSTTSDKPAPTTPKETAPTTPKEPAP-----TTP
                                                                                                                             TESSSAPVPTPS-SSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTE
                                                                                                                                                                                                        KAAAPNTPKEPAPTTPKEPAPTTPKEPAPTTPKETAPT-TPKGTAPTTLKEPAPTTPKKP
                                                                                                                                                                                                                                                          TAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPT-TPEEPTPTTPEEPAP-----TTP 535
                                                                                                                                                      APKELAPTTTKEPTSTTSDKPAPTTPKGTAPTTPKEPAP---TTPKEPAPT-TPKGTAPT
                                                                                                                                                                                                                                  SAPVT----SSTTESSSAP-VPTPSSSTTESSSAPVPTPSSSTTESSSAPVPTPSSSTTE
                                                                                                                                                                                                                                                                                    TPVTSSTTESSSAPVPTPSSSTTESSSAPVPT----PSSSTTESSSAPAPTPSSSTTESS 622
                                                                                                                                                                                                                                                                                                            AP----TTPKEPAPT-TPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKE
                                                                                                                                                                                                                                                                                                                                       SSAPVPTPS-SSTTESSSAPAPTPSSSTTESSSAPVTSSTTESSSAPVPTPSSSTTESSS
                                                                                                                                                                                         SSSAPVT---
                                                                                                                                                                                                                                                                                                                                                         EPAPTITKEPSPTTPKEPAPTITKSAPTTTKEPAP----TTTKSAPTTPKEPSPTTTKEP
                                                                                                                                                                                                                                                                                                                                                                                           ESSSAPVPTPSSSTTESSSAPVT---SSTTESSSAP-VPTPSSSTTESSSAPVTSSTTES
                                                                                                                                                                                                                                                                                                                                                                                                                    APTTPKEPTP----TTPKEPAPTTKEPAPTTPKEPAPTAPKKPAPTTPKEPAP----TTPK
                                                                                                                                                                                                                                                                                                                                                                                                                                              TESSSAPVTSS----TTESSSAPVTS---STTESSSAPVT----SSTTESSSAPVTSSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TSSTSESSTTTSSTSESSTSSSTTAPATPTTTSCTKEKPTPPTTTSCTKEKPTPPHHDTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DLSTGCNNYDNQGHSQTDFPGFYWNIDCDNNCGGTKSSTTTSSTSESSTTTSSTSESSTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                PVTSSTTESSSAPVPTPSSSTTESSSAPVTSSTTESSSAPVTSSTTESSSAPVPTPSSST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PALTTPKE-----PTPTTPKEPASTTPKEPTPTTTKSAPT----TTPKEPAPTTTKSAPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PCTKKKTTTSK-TCT---KKTTTPVPTPS-SSTTESSSAPV-----PTPSSSTTESSSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDLAPTSKVLAKPTPKAETT--TKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KVTTPDTSTTQHNKVSTSPKITTAKPINP------RPSLPPNSDTSKETSLTVNKETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210
817
874
1367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                               -SSTTESSSAPVT----SSTTESSSAPVPTPSSSTTESSSAPVPTPSSST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21
1367
1367
1367
817
874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11.78; 27.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                136110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polysaccharide degradation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         . . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLUCOAMYLASE (
SER/THR-RICH,
N-LINKED (GLC)
N-LINKED (GLC)
N-LINKED (GLC)
O MW; 91C00E2DB)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 798.5;
Pred. No. 5.9e
08; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JINKED (GLCNAC....
JINKED (GLCNAC....
91C00E2DBD61AA9D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5.9e-26;
hes 523;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S1/S2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .) (POI
.) (POI
CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels 177; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glycoprotein;
                                                 760
                                                                                                                             788
                                                                                                                                                                                                                                                                                                                481
                                                                                                                                                                                                                                                                                                                                       566
                                                                                                                                                                                                                                                                                                                                                                  426
                                                                                                                                                                                                                                                                                                                                                                                            507
                                                                                                                                                                                                                                                                                                                                                                                                                      370
                                                                                                                                                                                                                                                                                                                                                                                                                                              451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           50;
```

```
QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
STEPLECT STEPLECT OF THE PROPERTY OF THE PROPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       рь
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δÃ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SLP1_CLOTM
Q06852;
Q1-JUN-1994
Q1-JUN-1994
Q1-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLOTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2193
                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-NCIB 10682;
STRAIN-NCIB 10682;
REDLINE-93209931; PubMed=8458832;
RUJINO T., Beguin P., Aubert J.-P.;
Fujino T., Beguin P., Aubert J.-P.;
Fugino T., Beguin P., Aubert J.-P.;
Fujino T., Beguin P., B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Clostridium therman
namedia; Firmicutes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OLPB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            565
                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a copyright the Swiss Institute of Bioinformatics and the EMBL the European Bioinformatics Institute. There are no restrict use by non-profit institutions as long as its content is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTEIN 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           791
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     742
                                                                                                                                                       use by non-profit institutions as long modified and this statement is not removed. entities requires a license agreement (See or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=1515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        957
                                                                                              EMBL; X67506; CAA47841.1;
                                                                                                                                                                                                                                                                                                                                                                                                                        -!- SIMILARITY: CONTAINS 4 S-LAYER HOMOLOGY (SLH) DOMAINS
                                 InterPro; IPR001119; pfam; PF00395; SLH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                lostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     T-TPKETAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKEPTSTTSDKPAP-TTPKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TAPTTPKEPAPTTPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAPTTTKGPTSTTSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UN-1994 (Rel. 29, Created)
UN-1994 (Rel. 29, Last sequence update)
EB-1996 (Rel. 33, Last annotation updat-
SURFACE GLYCOPROTEIN 1 PRECURSOR (OUTE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R---TTP--ETTTAAPKMT---KETATTTEKTTESKITATTTOVTSTTTQDTTPFKITTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STPELSAEPTPKALENSPKEP-------GVPT-TKTPAATKPEMTTTAKDKTTERDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TQTPTTTPITTTTTTTPTPTPTGTQTPT-TTPITTTTTTTTTTPTPTGTQTPTT----TPIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KPAP-TTPKETAPTTPKEPAPTTPKKPAPTTPETPPPTTTSEVSTPTTTKEPTTIHKSPDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TTTTVTPTPTPTGTQTPT-TTPITTTTTVTPTPTPTGTQTPTTTPITTTTTTTTPTPTPTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QTPTTTPITTTTVTPTPTPTGTQTPTTTPITTTTTTTTPTPTPTGTQTPTTTPITTTTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKPKTMPRVRKPKTTPTPRKMTSTMPELNP---TSRIAEAMLQTTTR-PNQTPNSKLVEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KTTTLAPKVTTT-KKTITTTEIMNKPEETAKPKDRATNSKATTPKPQKPTKAP-KKPTST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TPTPTPTG---TQTP--TTTPITTTTTTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NPKSEDAGGAEGETPHMLLRPHVFMPEVTP 986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PS01072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          thermocellum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
          SLH_DOMAIN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus/Clostridium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            update)
                                                                                                                                                                                                       noved. Usage by and for commercial
(See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LAYER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Clostridiaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN
                                                                                                                                                                                                                                                                                                     restrictions
                                                                                                                                                                                                                                               and
                                                                                                                                                                                                                                                                                                                                                    a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B)
                                                                                                                                                                                                                                                   for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (S-LAYER
                                                                                                                                                                                                                                                                                                                                              outstation
                                                                                                                                                                                                                                                                              'n
                                                                                                                                                                                                                                                   commercia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         790
                                                                                                                                                                                                                                                                                 no
                                                                                                                                                                                                                                                                                                               on
                                                                                                                                                                                                                                                                                 way
                                                                                                                                                                                                                                                                                                               its
```

```
SET TO THE FIT THE FIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Дb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δõ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Вþ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       밁
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             В
                                                                                                                                                                                                                                                                                                      밁
                                                                                                                                                                                                                                                                                                                                                                          Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                    õ
                                                                                                                                                     멂
                                                                                                                                                                                                                     Ş
멍
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 312; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cell was SIGNAL CHAIN DOMAIN REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1095
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   193 VLAKPTP-KAETTTKGPALTTPKEPTPTTPKEPASTTPKEPTPTTIKSAPTTPKEPAPTT
    1397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  758
                                                                                                                                                                                                                                                                                                          1306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           367
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      585
                                                                                                                                                                                                                                                                                                                                                                                       759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VVIQPAPIKAASDEPIPTDTPSDEPTPS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEEPIPTDTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TKSAPTTPKEPAPTTTKEPAPTTPKEPA----PTTTKBPAPTTTKSAPTTPKEPAPTTPK 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TPSDEPTPSE----TPEEPIPTDTPSDEPTPSDEPTPSDEPTPSD-EPTPSDEPTPSETP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SDEPTSDEPTPSETPEEPIPTOTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TTPKEDA--PTTTKEPSPT-TPKEPAPTTTKS-APTTTKEPAPT---TTKSAPTTPKEPS: |: ||: ||: ||: || ||: || ||: || ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SDEPTPSDEPTPS--DEPTPS--DEPTP-SDEPTPTDTPSDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KPAPTTPKEPAPTTPKEPTPTTPKEPAPTTKEPAPT-TPKEPAPTAPKKPAPTTPKEPAP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PSDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSDEPTPSDEP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EEPIPTDTPSDEPTPSDEPTPSDEPTPSDEPTPSETPEBTPTDTPSDEPTPSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APT---
                                                                                                                                                                                                                                                                                                          SDEPTPSDEPTPSETPEEPIPTDTPSDEPTPSD----EPTPSDEPTPSDEPTPSDEPTPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PTPS--DEPTPSETPEEPIPTDTPSDEPTPSDEPT----PSDEPTP--SDEPTPSDEPTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAPTTPKKPAPT-TPETPPPTTSEVSTPTTTKEPTTIHKSPDESTPELSAEPTPKALENS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PKGTAPTTLKEPAPT-TPKKPAPKELAPTTTKGPTSTTSDKPAPT---TPKETAPTTPKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EPTPS--DEPTPSD-EPTPSDEPTPSETPEEPIPTDTPSDEPTPSDEPTPS--DEPTPS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EEPAPTTPKAAAPNTPKEPAPTTPKEPA--PTTPKEPAPT-TPKETAPTTPKGTAPTTLK
SGGSGGGGGGGGGTVPTSPTPTPTSKPTSTPAP--
                                                                                                                                                                                                                         SKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEETAKPKDR
                                                                                                                                                                                                                                                                                                                                                                              PKEPGVPTTKTPAAT - - KPEMTTTAKDKTTERDLRTTPETTTAAPKMTKETATTTEKTTE
                                                                            ATNSKAT-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -DEPTPSDEPTPSETPEEPI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PTTTKEPAPT-TPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEP 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1378
1453
1495
1566
1626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S-layer;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TPKETAPTTPKKLTPTTPEKLAPTTPEKPAPT----TPEELAPTTPEEPTPT-TP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29
36
36
207
409
607
771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1449
1494
1565
1625
1646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1664
763
191
363
363
565
763
1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Signal;
28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13.9%;
31.0%;
                                                                            -TPKPQKPTKAP-KKPTSTKKPKTMPRVRKPKTTPTPRKM----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         178194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPROXIMATE TANDEM REP
T-P-S-D-E-P-
GIX/PRO/SER/THR-RICH.
SLH 1 (INCOMPLETE).
SLH 2.
SLH 3.
SLH 4 (INCOMPLETE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Repeat.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELL SURFACE GLYCOPROTEIN 1.
4 x 156 AA APPROXIMATE REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 (INCOMPLETE).
5F396695BA9FE74B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                950; DB 1;
No. 5.1e-32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PTDTPSDEPTPSDEPTPSD-EPTPSDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----DEPTPS---DEPTPSDEPTPSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TANDEM REPEATS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361;
                    -TEIEEPTPSDVPGAIGGEHRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1664;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       230;
                                                                                                                                                                             PTSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        366
                                                                                                                                                                                                                                                                                                                                                                                                           816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1154
                                                                                                                                                                                                                                                          876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1094
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    758
                                                                                                    921
                                                                                                                                                                             1396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               55
```

N

```
THE PROPERTY OF THE PROPERTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; L21998; AAB95295.1; -.
EMBL; M74027; AAA59975.1; -.
EMBL; M94131; AAA59163.1; -.
EMBL; M94132; AAA59163.1; -.
EMBL; M94132; AAA59164.1; -.
MIN; 158370; -.
InterPro; IPR000359; Cys_knot.
InterPro; IPR002400; GF_cysknot.
InterPro; IPR001007; VWFC.
InterPro; IPR001007; VWFC.
InterPro; IPR001007; VWFC.
InterPro; IPR00140; VWGC; 1.
Pfam; PF000007; Cys_knot; 1.
Pfam; PF000094; VWG; 4.
PRINTS; PR00438; GFCYSKNOT.
SWART; SW00211; VWC. def; 2.
SWART; SW00211; VWC. def; 2.
SWART; SW00216; VWD; 4.
PROSITE; PS01208; CTCK_1; 1.
PROSITE; PS01208; VWFC; 2.
GLYCOPTOTENTIA.
GLYCOPTOTENTIA.
        Glycoprotein;
SIGNAL
CHAIN
CHAIN
CHAIN
CHAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
             51747
11416
11416
11416
11416
11416
11417
11417
11417
11417
11417
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
11517
     POTENTIAL,
MUCIN 2.
APPROXIMATE REPEATS.
(POTENTIAL)
     В
```

```
OM protein - protein search, using sw model
                                                                                 GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on: April 26, 2002, 16:32:21; Search time 62.75 Seconds (without alignments) 742.061 Million cell updates/sec

Title:
Perfect score:
Sequence: AA3
6814
1 MAWKTLPIYLLLLSVFVIQ.....ARAITTRSGQTLSKVWYNCP 1270

Searched:

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 100059

100059 seqs, 36664827 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_39:*

Database :

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

28 29 30 31 33	20 21 23 24 25 26	1 1 2 2 3 3 3 3 3 4 4 4 5 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6	Result
439.5 432.5 432.5 432.4 430.5 424.5	472 471.5 468.5 467 457 454 454 448.5	1171 950 798.5 651.5 556.5 555.5 53.0 530.5 530.5 530.5 530.5 498.5 498.5 493.5 488.5	Score
	5.5.7.999 5.6.7.999	113.9 113.9 113.9 114.0 115.0	% Query Match L
1083 439 3164 3421 2715 1125	2142 5376 634 817 1161 797 670	5179 11664 11367 1255 875 2700 620 1087 1162 862 862 1970 1970 1970 872 664 872 664 872 664 872 665 872 666 873 874 875 875 875 875 875 875 875 875 875 875	Length DB
1 TEGU_HSVEB 1 TEGU_HSV11 1 TEGU_HSV11 1 TEGU_HSVEB 1 TEGU_HSVEB 1 TEGU_HSVEB 1 TEX2_HUMAN 1 MAP4_MOUSE	1 BAT2_HUMAN 2AN_MOUSE 1 HWP1_CANAL 1 VRP1_YEAST 1 YJ9P_YEAST 1 YJ9P_YEAST 1 YGLX_HSVEB 1 VG50_HSV11 1 VG50_HSV11 1 FP1 MYTGA	I MUCZ_HUMAN I SLP1_CLOTM I AMYH YEAST MUC1_HUMAN FP1_MYTED I ZAN_HUMAN EXTN.TOBAC I NFH_MOUSE I TCNA_TRYCR I CPN_DROME I FP1_MYTCO I MUC1_SCNLA RPB1_HUMAN I RPB1_HUMAN I RPB1_CRIGR SSP2_PLAYO I NFH_SMU SSP2_PLAYO I NFH_CRIGR SSP2_PLAYO I NFH_CRIGR SSP2_PLAYO I NFH_CRIGR I RPB1_CRIGR SSP2_PLAYO I NFH_CRIGR I RPB1_CRIGR I RPB1_CRIGR I RPB1_CRIGR I NFH_STMU I NFH_LMOUSE I NFH_STMU	BID
MAN LA V11 VEB MAN USE	MAN SE SE AST AST VEB	AAN AST AST AN AN AN BAC BAC CC CC CC CC MAN)
			1
000268 P17437 P10220 P18955 P28955 Q9umn6 P27546	P48634 O88799 P46593 P37370 P47179 P28968 Q00130 O27409	7304080408044808	Description
homo sapien mus musculu	homo sapien mus musculu candida alb saccharomyc saccharomyc equine herp ictalurid h mytilus gal		otion

5	44	43	42	41	40	39	3 8	37	36	ω 5	34
407.5	410	410	411.5	419	419.5	420	421	422	422.5	424	424
6.0	6.0	6.0	6.0	6.1	6.2	6.2	6.2	6.2	6.2	6.2	6.2
3256	2517	1183	907	1185	2476	1411	1229	2774	1794	1251	307
1	ш	Н	ш	щ	\vdash	Н	₩	Н	سا	\vdash	1
KI67_HUMAN	NCR2_HUMAN	DRPL_RAT	VGP3_EBV	DRPL_HUMAN	ZAN_PIG	TCOF_HUMAN	N121_HUMAN	MAPA_RAT	YAV1_SCHPO	YQU3_CAEEL	SGS3_DROME
P46013	Q9y618	P54258	P03200	P54259	Q28983	Q13428	Q9y2n3	P34926	Q10172		P02840
homo saj	h nuclea	rattus n	epstein-	homo sap	sus scroi	homo sapi	homo sapien	rattus no	schizosac		drosophila

ALIGNMENTS

Page 13